Radboud University

Nijmegen, the Netherlands

Investigating the acceptance of MyST, an innovative tool for the improvement of spoken academic English, with the UTAUT2 research model

Angela Pantazi
s1029860
5/7/2019
Master's Thesis (LET-TWM400)
Primary Supervisor: Helmer Strik
Secondary Supervisor: Catia Cucchiarini

Acknowledgments

Foremost, I would like to express my sincere gratitude to my supervisor Dr. Helmer Strik, for his guidance, motivation, and detailed comments during the time of research and writing of this thesis. Besides my advisor, I would like to thank my fellows Camilla Giannini and Anna Ovchinnikova for our cooperation; this study could not be conducted without their contribution, as they shared with me the collected data from the students of Utrecht and Nijmegen University.

I am also extremely grateful to my family, Giorgos Pantazis, Jenny Tsarmpopoulou and Panagiotis Pantazis for the essential support to new steps in my life, as well as for caring and sacrifices for educating and preparing me for my future. Last but not least, I would like to thank one of the most important people in my life, Ilias Tsafaras, who inspire, motivate and encourage, me with his optimism and devotion, to get out of my comfort zone and try new exciting things in my career.

1. Contents

2. Abstract

Many universities worldwide are offering courses of EAP implementing CALL systems in teaching and learning. However, CALL systems with ASR-technology that develop communication skills are still limited, and they do not provide academic English content. *My Speech Trainer* is one ASR-based system intended to help students improve their general academic English speaking skills, as well as their subject-specific language knowledge. This study applied an extended version of UTAUT 2 model aims to investigate behavioral, organizational, and individual factors that affect student's acceptance of MyST, contributing to its development and its successful implementation in higher education. PLS-SEM technique and descriptive statistics were employed to analyze the data collected from students of Nijmegen and Utrecht University (n= 84), as well as a multi-group analysis was performed to investigate the moderating effects of organizational factors and faculty in behavioral relationships. Results show that SI was one of the main predictors of students' intention to use the system, following by At. Organizational factors and faculty did not moderate SI, while both of them moderate At. In general, the results show that participants were generally positive about *My Speech Trainer*. Finally, this research expanded the application of the UTAUT2 model in ASR-based CALL systems and extended it, adding the construct of At and the moderating factors.

Keywords: MyST, ASR-based CALL, EAP, course-specific content, UTAUT2, PLS-SEM, multi-group analysis, moderating effects, attitude

3. Introduction

Due to the rapid growth of technology, more and more universities are investing in electronic learning (e-learning) and mobile learning (m-learning) systems to support student's knowledge and performance. Computers and related technologies are considered important media in learning and teaching in universities (Adair-Hauck, Willingham-McLain, & Youngs, 2000; Tarhini, Hone, & Liu, 2013), including not only within disciplines' related courses but also in language learning courses (Cavus & Ibrahim, 2009; Lowenthal, 2010; Nassuora, 2012). Considering the student's need of learning and using academic English as well as English related to their disciplines, several international universities are offering courses of English for Academic Purposes (EAP) implementing new technologies and Computer Assisted Language Learning Systems (CALL).

Although students may have a proficient level in English to be accepted by universities, they still face many challenges in communication, pronunciation, and vocabulary. Therefore, it is crucial for them to develop further the specific language and terminology required in an academic environment as well as to do additional practice for improving their English oral skills. However, practicing speaking skills is time-consuming and requires intensive training, access in authentic pronunciation, and focused feedback, which cannot always be provided to a sufficient extent in class. Most of the times, University's curricula focus on topics of the specific academic field, and they do not include speaking practice. Furthermore, many students hesitate to practice their speaking skills, as they are afraid or embarrassed from practicing speaking in public (Chiu, Liou, & Yeh, 2007; Elimat & AbuSeileek, 2014).

Therefore, the development of Computer Assisted Language Learning (CALL) systems that employ Automatic Speech Recognition technology (ASR), usually called CAPT (Computer Assisted Pronunciation Training) (Neri, Cucchiarini, & Strik, 2003), could offer tailor-made courses for the enhancement of students' oral skills and subject-specific language knowledge (Golonka, Bowles, Frank, Richardson, & Freynik, 2014; Neri, Cucchiarini, & Strik, 2001; Neri, Cucchiarini, & Strik, 2006; Zhao, 2003). The implementation of ASR in CALL systems offer many advantages to EFL (English as a foreign language) learners; it provides additional practice in a low-anxiety environment, explicit and individualized feedback on speaking skills and enhancement of learners' independence for language learning (Cucchiarini & Strik, 2018; Elimat & AbuSeileek, 2014; Neri, Cucchiarini, & Strik, 2003; Neri, Mich, Gerosa, & Giuliani, 2008).

As regards the context of CALL and ASR, technology alone cannot lead to improved acquisition of English by EFL learners; for successful implementation and outcomes, it is necessary for course designers and instructors to have a thorough understanding of learners' attitudes and acceptability towards ASR-based systems (Hsu, 2016; Mah & Er, 2009). Although, there is an extensive literature for the acceptance of CALL systems by learners (Afshari, Ghavifekr, Siraj, & Jing, 2013; Soleimani, Ismail, & Mustaffa, 2014; Wiebe & Kabata, 2010), little empirical research has been carried out in the field of CAPT systems investigating the relationship of behavioral, institutional and individual factors with learners' intention to use ASR-based tools. Most of the research conducted so far focuses on the function and performance of ASR technology itself (Foote & McDonough, 2017; Golonka et al., 2014; Savvani, 2018), and less on the real students' experience with the tool (Cucchiarini & Strik, 2018; Hsu, 2016; Neri, Mich, Gerosa, & Giuliani, 2008). Furthermore, although previous studies have already focused on the students' perception in a foreign language teaching setting (Ayres, 2002; McCrocklin, 2019; Wiebe & Kabata, 2010), many concerns remain on the use of ASR-based CALL systems in the academic context and specifically in learning English for Academic Purposes (EAP).

In light of the above, the present study aims to investigate the student's acceptance of an ASR-based CALL system used for the improvement of spoken academic English. As far as we know, ASR-based CALL systems have not yet been developed for the specific purpose of improving oral academic English, and previous studies have not investigated the role of behavioral constructs and other factors in the acceptance of CAPT systems. Therefore, this study expects to address these gaps in research by presenting a prototype application called *My*

Speech Trainer (MyST) and by investigating its perception by students applying a technology acceptance model, the UTAUT2 model.

Previous literature shows that the success of new technology is often attributed to several behavioral, individual, social, organizational, and cultural factors (Browne et al., 2006; Schepers & Wetzels, 2007; Tarhini et al., 2016). These factors are essential for the successful adoption of a system, as research indicates that the acceptance of tools by students vary across population groups, societies, and cultures (Dečman, 2015; El-Masri & Tarhini, A., 2017; Teo & Noyes, 2014; Venkatesh & Zhang 2010). Therefore, students' acceptance of the MyST ought to be taken into account focusing on behavioral, organizational, and individual factors. The behavioral factors are investigated by applying to the whole sample a revised version of the Unified Theory of Acceptance and Use of Technology, the UTAUT2 model, which was extended through the inclusion of attitude (Venkatesh, Thong & Xu, 2012; Venkatesh, Morris, Davis & Davis, 2003). Furthermore, the UTAUT2 model is applied to different population groups for examining the role of organizational factors and student's faculty in acceptance of the MyST. The whole sample was divided applying two criteria; the first is the context of the use of MyST, which is identified with the university that the students were studying (Utrecht University – Nijmegen University) and the second is the faculty of students (Languages and Communication – Other studies). A multi-group analysis was performed for each category of groups to be investigated the relationship of constructs depending on the group.

Therefore, the first goal of this research is to investigate the relationship between the behavioral constructs of Performance Expectancy (PE), Effort Expectancy (EE), Attitude (At), Habit (Ht), Social Influence (SI) and Facilitating Conditions (FC) with user's Behavioral Intention (BI) to use the MyST expecting to predict the future use of the system. Furthermore, extending this goal, the present study intends to explore if there are differences in the relationships mentioned above, considering the moderating effects of organizational factors and faculty. Therefore, the present study was designed to answer the following research questions:

1. How do the behavioral constructs affect the behavioral intention (BI) of users to use the MyST?
2. How do organizational factors and faculty affect the relation between the behavioral constructs and behavioral intention (BI) of users to use the MyST?

The remainder of this study is structured as follows. It begins with the theoretical background of the research related to the MyST, the UTAUT model, and the formulation of research hypotheses. Then, it presents the methodology that describes the participants, the procedure for the collection of data and the statistical method of analysis. After that, the data analysis and results are described. Finally, the study ends with a discussion of research findings, implications for future research and concluding remarks.

4. Theoretical Background

In this section, background information is provided about the MyST tool, the 'Unified Theory of Acceptance and Use of Technology' (UTAUT), the research model of the study and the use of technology acceptance models in CALL.

4.1 My Speech Trainer

My Speech Trainer application was developed using the NovoLearning Editor, and with the NovoLearning Player it can be used on different devices: smartphone, tablet and PC (https://www.novolearning.com/). Learners could operate the system through mouse and keyboard, but they could also speak and get individual and immediate feedback on their spoken utterances on various aspects, such as pronunciation, lexicon, syntax, grammar, and vocabulary. These functions are provided due to the ASR technology, the Novo-CALL, which is an intelligent software that recognizes and checks speech. The novelty of the system relies on the combination of ASR technology with the context of academic language. The practice offered in MyST addresses general academic English, while content for specific courses could be easily added through an editor. Besides, other users such as teachers can also do this independently, as they can add exercises which also include feedback on spoken utterances. In other words, MyST aims to improve oral communication skills, particularly in academic English. This includes enhancing pronunciation skills, using collocations, and knowledge of language customs, e.g., register through the ASR system that evaluates the pronunciation quality. Some essential features of the system are that it provides error detection and error diagnosis, it locates the errors in the utterance, it identifies the specific type of error, and it suggests to the learner how to improve it. Finally, MyST has a feedback presentation in which presents the overall score as a percentage in a pie-chart, visualizing the progress of the learners. These functions are so essential that they can make a CAPT system the ideal system, as stated by Neri et al. (2003, 1165).

4.2 The Unified Theory of Acceptance and Use of Technology (UTAUT model)

With the growing reliance on technology, technology acceptance models[1] were developed to identify behavioral factors affecting people's intentions to use new technologies, and thereby to promote acceptance of such technologies. One recent related model that is widely used in Information and Communications Technology (ICT) research, but not yet as widely in CALL, is the Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh et al., 2003). The model was created through extensive review as well as empirically compared eight acceptance models which are theory of reasoned action (TRA), technology acceptance model (TAM), motivational model (MM), theory of planned behaviour (TPB), model of PC utilization (MPCU), innovation diffusion theory (IDT), combined TAM and TPB (C TAM-TPB) and social cognitive theory (SCT). Therefore, the comparison of acceptance models resulted in the formulation of the UTAUT model. It has four direct determinants of behavioral intention (BI) and actual use: performance expectancy (PE), effort expectancy (EE), social influence (SI) and facilitating conditions (FC). It also has four moderators that are found to mediate the effects of the four key determinants on usage intention and behavior: the age, the gender, the voluntariness of use, and the experience. The UTAUT model was able to explain 70 percent of the variance in ICT use and to do so with greater clarity than the other models (Venkatesh et al., 2003). In 2012, Venkatesh, Thong & Xu (2012) proposed the UTAUT2 model with an improved variance that included three new constructs to the UTAUT model: Hedonic Motivation (HM), Price Value (PV) and Habit (Ht) (Figure 1).

Both versions of UTAUT have been widely used in research to examine users' acceptance of various technologies in different fields. Williams, Rana & Dwivedi (2015) support with their review of 314 studies that the model has been used for tools related to the

[1] For various technology acceptance models see Tan, 2013.

government, the digital learning, the education, the social media as well as for communication tools such as mobile phones and online banking for specific business and office technologies. However, UTAUT has been criticized for showing bias across different settings (Dwivedi, Rana, Chen, & Williams, 2011; Teo, 2015). Consequently, Venkatesh et al. (2012) propose the testing of UTAUT2 in different contexts to enhance its applicability, as factors that affect the adoption of new technologies differ depending on the context, target users and technology. Considering the simplicity and the robustness of UTAUT2 to analyze behavioral intentions to the reception of new technologies in diverse situations (Venkatesh & Zhang, 2010), it is adopted as the framework and the research model of this study for explaining the acceptance of MyST.

4.3 Research Model and Theory

Although the study is based on the revised version of UTAUT, the UTAUT2 model (Venkatesh et al.) (Figure 1), some alterations have been made in our research model to be adapted adequately in the settings of the study. First of all, as many other studies (El-Masri & Tarhini, 2017; Morosan & DeFranco, 2016), this study did not consider the moderating effect of gender, age, and experience. The participants are university students, and they have similar age and experience. In addition, the unequal sample size between men and women was not suitable for applying the moderating factor of gender. However, the present study investigates the moderating effect of organizational factors and of faculty.

Besides, as MyST is a free technological resource, and it does not incur financial costs, the price value was removed from this model. On the other hand, it was added the construct of attitude (At), as recommended by the study of Dwivedi, Rana, Jeyaraj, Clement & Williams (2017), which supports that the attitude has a significant role in the model. Finally, our research model considers as dependent variable only the behavioral intention (BI) and not the actual use of the tool, as the way of data collection for the participants from Nijmegen University did not allow the creation of user's personal accounts, and thus the estimation of real time of use. Therefore, our research model includes the relationship between the constructs of Performance Expectancy (PE), Effort Expectancy (EE), Hedonic Motivation (HM) Attitude (At), Habit (Ht), Social Influence (SI) and Facilitating Conditions (FC) with user's Behavioral Intention (BI) to use the MyST (Figure 1).

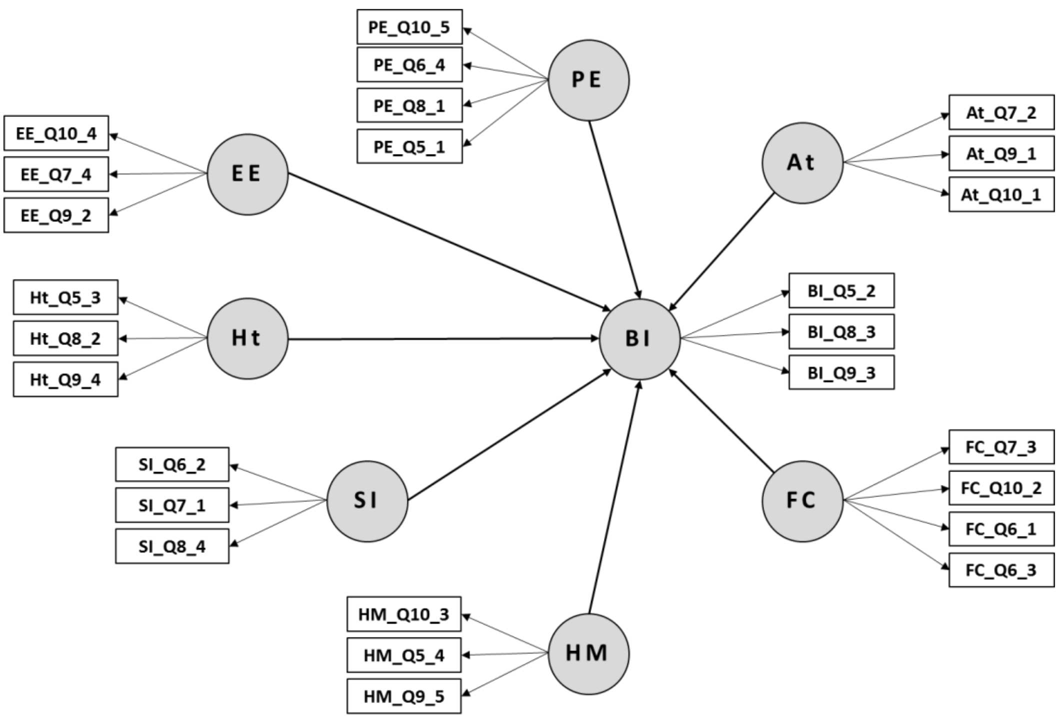

Fig. 1 First model: constructs (circles) and indicators (boxes).

Many studies use the UTAUT model in organizational settings, but relatively few studies used it in the context of CALL. As the relationship between the constructs and the intention to use vary per tool, target audience and context, the definition of constructs and their relations based on previous literature are essential for the formulation of the hypotheses of our study.

Attitude

Venkatesh et al. (2003) defined the attitude toward using technology as “an individual's overall affective reaction to using a system”. However, the UTAUT model does not include attitude because the statistical analysis did not show any significant relationship between At and BI (Venkatesh et al., 2003). On the contrary, Dwivedi, Rana, Jeyaraj, Clement, & Williams (2017) proposed an alternative model that implements the attitude (At) to the UTAUT model. In their meta-analysis of 162 studies, they observed that attitude affected BI and actual use, and hence, they included into the model. Based on previous literature (Dwivedi et al., 2017), this study implemented the At to the UTAUT2 model extending it. Therefore, the following hypothesis is proposed:

H1: Attitude will have a positive and significant influence on students’ intention to use the MyST.

Facilitating Conditions

FC defined as user’s perceptions of the resources and support available to perform a behavior (Venkatesh et al., 2012). Explicitly, resource determinants such as time for training and system’s compatibility are essential features of facilitating conditions (Lu, Liu, Yu, & Wang, 2008). It is reported that users will not have the willingness to use the system if they do not have adequate time for training, or there are issues related to the system's infrastructure. Therefore, FC may influence a person's perception of his performance expectancy as well as the level of difficulty of new technology. Previous studies have found a significant influence of FC on BI to use a technology (Teo, 2009; Teo, Lee, & Chai, 2008; Khechine et al., 2014). However, in most of these studies (Teo, 2009; Teo, Lee, & Chai, 2008), FC has indirect effects on the intention to use technology. Venkatesh et al. (2003) suggested that the issue relating to FC (e.g., support infrastructure) is captured within the EE, which refers to the ease with which technology can be used. As such, we may consider that when both PE and EE are present, FC is not significant in predicting BI of use technology (Teo & Noyes, 2014). From the above discussion, the following hypothesis is proposed:

H2: FC *will not have* a significant influence on students’ intention to use the MyST.

Social Influence

SI is “the degree to which an individual perceives that important others believe he or she should use a new information system” (Venkatesh et al., 2003). The standpoint of the user's friends, families, professors, and seniors affect user behavior (Lopez-Nicolas, Molina-Castillo, & Bouwman, 2008). Previous literature shows that social influence is weighty in changing a person’s intention to use technology (Pynoo, 2011; Teo, 2009). In this study, social influence refers to those important people for students such as university professors and friends, who think that students shall or shall not use the MyST. Thus, the following hypothesis is aroused:

H3: Social influence will have a positive and significant influence on students intention to use the MyST.

Habit

Kim and Malhotra regarded habit as an equal with automaticity. Limayem, Hirt, & Cheung (2007) defined habit as "the extent to which an individual tends to perform behaviors automatically”. Venkatesh et al. (2012) defined habit as a perceptual construct that reflects the results of an individual’s experiences. Despite the different definitions of habit, they share the same idea that the feedback from previous experiences will affect behavioral intention. Nair, Ali, & Leong (2015) investigated the factors affecting students’ acceptance and usage of a lecture capture system, the ReWIND, employing the UTAUT2 model, and they found that Ht is a significant and positive predictor of BI. Furthermore, Kang, Liew, Lim, Jang, & Lee (2015) found that habit affects significantly behavioral intention to use m-learning among Korean university students. From the above discussion, hypothesis was proposed as follows.

H4: Habit will have a positive and significant influence on students intention to use the MyST.

Effort Expectancy

EE is defined as the degree of ease associated with the use of technology (Venkatesh et al., 2003). Based on the UTAUT, use of MyST among educational users will depend on whether or not the technology is easy to use. Previous literature support that EE is a significant predictor of BI; Hsu (2016) investigated the use of the Moodle platform for Computer-Mediated Communication. Collecting questionnaires from 47 students in Taiwan and running a regression analysis, they found that EE has a significant and positive influence on BI. Besides, Liu & Huang (2015) researched the use of Google Docs for translation practice with 27 students and found EE to be a significant and positive predictor of BI. Thus, this study tests the following hypotheses:

H5: Effort expectancy will have a positive and significant influence on students' intention to use the MyST.

Performance expectancy

Venkatesh et al. (2003) defined PE as 'the degree to which an individual believes that using the system will help him or her to attain gains in job performance'. In an educational context, PE means that users will find the technology useful, as it facilitates their progress in learning activities. Previous literature support that PE is a strong predictor of BI; Tan (2013) investigated the use of websites for learning English in Taiwan, collecting 176 responses to a UTAUT-questionnaire. The study reported PE as one of the strongest predictors. Furthermore, Khechine, Lakhal, Pascot, & Bytha (2014) collected 114 UTAUT-questionnaires to investigate acceptance and use of webinar platform for various disciplines, including language learning, and found that PE was one significant predictor. From the above discussion, hypothesis was proposed as follows:

H6: Performance expectancy will significantly and positively influence students' intention to use the MyST.

Hedonic Motivation

HM is defined as the degree of fun or pleasure from using a technology (Venkatesh et al., 2012). In the UTAUT model, the extrinsic motivation for the technology use is defined as performance expectancy (Venkatesh et al., 2003). As regards the motivation theory, apart from extrinsic motivation, intrinsic or hedonic motivation also plays an important role in the decision for technology use. Hedonic motivation is a key predictor of technology acceptance in many educational settings (Hong, Thong, & Tam, 2006; Yang, 2013). Therefore, the following hypothesis is proposed:

H7: Hedonic Motivation will have a positive and significant influence on students' intention to use the MyST.

Moderating effects

As already mentioned, the present study takes into account as moderating effects organizational factors related to the use of MyST and the faculty of students that used the MyST. As regards the organizational factors, they include the voluntariness, and the nature of the task and the professions in the working settings (Sun & Zhang, 2006). In our research, which has conducted in an educational academic context, the voluntariness is defined as the extent to which students perceive the adoption decision of the MyST to be non-mandatory (Venkatesh & Davis, 2000), while the task mentions to the routine and non-routine of using the tool (Sun & Zhang, 2006). Furthermore, in the academic context of our study, the task is related to the course-specific content that has been added to MyST and has been implemented as teaching material on an academic course. More specifically, the sample was separated into two subsamples depending on the organizational factors. The first subsample from the University of Nijmegen used the MyST voluntarily for a short period and did activities with general content, as the participants used the MyST in the framework of promotion of the tool. On the other hand, the second subsample from the University of Utrecht used the MyST semi-voluntarily for an extended period, as it was implemented in the syllabus of academic courses. Therefore, students did specific-course content activities. Previous research trying to extend the UTAUT model has tested many different moderating factors such as voluntariness (Venkatesh et al., 2012), previous education (Schepers & Wetzels, 2007), gender (Khechine et al., 2014) and culture (Tarhini et al., 2016). However, little research exists that takes into consideration the organizational factors in an academic context (Sun & Zhang, 2006).

Furthermore, for investigating the moderated effect of faculty, the sample was separated again in two subsamples; the participants of the first subsample were studying Languages and Communication and the participants of the second subsample were studying in other disciplines. To the best of our knowledge so far, the research has not yet examined the faculty of students as a moderating effect in the UTAUT model. Previous literature applies the UTAUT model in different faculties such as in Social Studies and Business Administration Faculties (Attuquayefio & Addo, 2014), but it does not compare the relationships between behavioral constructs considering the different disciplines.

In this study, we examine the validity of the UTAUT2 model across the whole sample, and at the same time, we test the model across the subsamples implementing the organizational factors and faculty. The comparison of different groups could lead to a broader view and a better explanation of MyST adoption, as the acceptance of tool is investigated focusing on different groups of the population.

Furthermore, apart from examining the proposed hypotheses, we also investigate the potential organizational and faculty differences associated with the research model. Since there is a lack of theoretical support of evidence from existing literature, we do not propose any hypothesis on these differences. We conduct an exploratory analysis to investigate the existence of faculty and organizational differences by testing the research model on two subsamples. This way of investigating differences in technology acceptance models has been adopted in recent literature (Padilla-Meléndez, Aguila-Obra, & Garrido-Moreno, 2013).

In general, the present study attempts to bridge the gap between technology and language learning, to evaluate the acceptance of a tool with ASR technology and to tackle the problem of inconsistent methodologies in CALL research. Furthermore, it aims to help stakeholders to gain a deeper understanding about the factors that affect the students' decisions to adopt MyST, and thus, they are able to formulate adequate strategies that encourage the adoption of MyST. Besides, in the research field of technology acceptance, this study contributes to the literature related to theories and models of technology adoption that has recommended an expansion of the models in new contexts (e.g., Venkatesh & Zhang, 2010). Explicitly, it expands the UTAUT2 model through the inclusion of construct of attitude, and through the implementation of organizational factors and faculty as moderating effects. Finally, it contributes to the generalizability and applicability of the UTAUT2 in the new context of CAPT systems.

5. Methodology

5.1 Research Instrument

The questionnaire was created in Qualtrics (2005) and it was accessed through an internet link. It consisted of three parts. The first part included a content form and contact data (email). The second part consisted of 26 UTAUT items, that represent the predictor constructs of PE (4 items), EE (3 items), SI (3 items), FC (4 items), At (3 items), Ht (3 items), HM (3 items) and the target variable of BI (3 items) (Appendix 2). Survey items have been drawn from the literature (Dwivedi et al., 2017; Venkatesh et al., 2003; Venkatesh et al., 2012) where they were quoted to be reliable and valid to measure constructs of the phenomena that they intend to represent, and they were modified to fit our context of MyST. The participants were asked to rate the statements on a 7-point Likert scale format (1- Strongly disagree, 2 - Disagree, 3 - Somewhat disagree, 4 -Neither agree nor disagree, 5- Somewhat agree, 6 - Agree, 7 - Strongly agree). The final part of the questionnaire listed questions on personal (gender, age, nationality), linguistic (self-assessment of English knowledge) and educational backgrounds (university, degree level, program of enrolment). On average, each participant took approximately 20 min. to complete the questionnaire.

5.2 Participants

Nijmegen University

Participants were recruited in several ways at Radboud University in Nijmegen: through lecturers, on the spot at the Radboud University campus and through personal social networks. Most responses were collected by approaching students at the campus (two canteens on the Radboud University campus). There were in total 48 participants, and the vast majority (n= 41, 86%) was from 18 to 24 years old. They were students that enrolled in bachelor (n=40, 83%) and master's (n=8, 17%) degrees from different faculties, such as Social Sciences faculty (n=17, 35%), Arts (n=15, 31%), Applied Sciences (n=10, 21%) and other faculties (n=6, 13%). The students who agreed were instructed to complete voluntarily minimally 2-3 exercises in MyST. Once the students got an impression of MyST functionality and content, they filled out the questionnaire. The participants, who were approached at the campus, were offered a chocolate bar as a reward, and they were told that five participants would receive a gift coupon of 10 Euro each. After data collection was completed, we randomly selected these five participants.

Utrecht University

The participants from Utrecht University were 36 students who were following a Semantics and Pragmatics course, and a Phonetics course as part of their bachelor's degree. The age of participants ranged from 18-25 with average age 20 years old. The Semantics and Pragmatics course took place from February until April 2019, and the Phonetics course from April until June 2019. In the course of Semantics and Pragmatics, there were 31 students. In this case, only the data from the 11 students, who completed the questionnaire, were analyzed. Many students did not use the MyST or dropped out from the course before the end, and thus, they did not respond to the questionnaire. Even though all the students were enrolled in a bachelor course, there was a great variety regarding the faculties. Most of the students (i.e., 26) were enrolled in different courses at the Faculty of Humanities, 8 at the Faculty of Science, and 2 at the Faculty of Social and Behavioural Science. In the course of Phonetics, there were enrolled 51 students. Not all the students did the activities in MyST. The students who finished at least one activity were 30, but among these, 25 filled out the UTAUT questionnaire.

Professors presented MyST the first day of class. Using the NovoLearning editor, we added specific content to MyST for each of these two courses. Each week, during the preparation for a class, the students could use MyST containing content for the class of that week. The students were told to use MyST as an extra, non-graded, assignment, but the instruction was included also in the courses' syllabus, which means that the use of MyST was in a semi-voluntary context. After using MyST, the participants were given the questionnaire. In total, 36 students completed the questionnaire, 11 from the Semantics and Pragmatics course, and 25 from the Phonetics course.

Overall sample

The target sample for this survey includes 84 participants and it is ensured sufficient variation (Kock & Lynn, 2012), as it came from Dutch and international users of MyST, enrolled in masters or undergraduate programs in several disciplines at the University of Nijmegen and Utrecht. The average age of participants was 20, and the majority of the participants were female (71,4%). The vast majority of them were enrolled in bachelor's degrees (90,5%). Considering the fields of studies, 36 of students (44,4 %) were studying languages and/or communication, and 45 of students (53,6 %) were enrolled in other disciplines. Finally, 32,1 % of the participants self-rated themselves as of intermediate proficiency level (Table 1).

Table 1 Demographic data.

Characteristics	Frequency	Percent	Valid Percent	Cumulative Percent
University				
Nijmegen	48	57.1	57.1	57.1
Utrecht	36	42.9	42.9	100
Faculty				
Valid: Languages/Communication	36	42.9	44.4	44.4
Valid: Other Faculty	45	53.6	55.6	100.0
Valid: Total	81	96.4	100.0	
Missing: -1.00	2	2.4		
Missing: System	1	1.2		
Missing: Total	3	3.6		
Total	84	100.0		
Spoken English Proficiency				
B1	10	11.9	11.9	11.9
B2	17	20.2	20.2	32.1
C1	38	45.2	45.2	77.4
C2	19	22.6	22.6	10
Degree				
Bachelor's	76	90.5	90.5	90.5
Master's	8	9.5	9.5	100
Gender				
Female	60	71.4	71.4	71.4
Male	24	28.6	28.6	100
Nationality				
Dutch	69	82.1	82.1	82.1
International	15	17.9	17.9	100
Age				
18	7	8.3	8.3	8.3
19	19	22.6	22.6	31
20	17	20.2	20.2	51.2
21	17	20.2	20.2	71.4
22	7	8.3	8.3	79.8
23	7	8.3	8.3	88.1
24	2	2.4	2.4	90.5
25	8	9.5	9.5	100

5.3 Method of Analysis

Consistent with previous empirical research in technology acceptance (e.g., Venkatesh & Bala, 2008), the current study adopted a quantitative approach to test the proposed model. The data were analyzed using IBM SPSS Statistics 22, mainly for the descriptive statistics, and partial least squares structural equation modeling (PLS-SEM) for examining the relationships among variables within the proposed conceptual model. Despite criticism, PLS is a well-established technique for estimating path coefficients in structural models, and it has the ability to model latent constructs under conditions of non-normality and small-to-medium sample sizes (Ali, Hussain, & Ragavan, 2014; Hair, Hult, Ringle, & Sarstedt, 2014). Furthermore, SEM approach was chosen over the usual regression analysis, because it permits simultaneous analysis of the relationships among variables and errors for each variable to be independently estimated (Hair et al., 2014). In overall, PLS-SEM was performed and found to be suitable in this study, as it is adequate for small sample size, it focuses on prediction (Henseler, Ringle, & Sinkovics, 2009) and it assumes all measured variance useful for estimating interaction and main effects (Chin, Marcolin, & Newsted, 2003).

The software SmartPLS 3 was used for the application of PLS-SEM. The measurement and structural model were assessed only for the overall sample and not for the group sample, as this assessment is not needed when the model for the group samples presents measurement invariance (see section 6.6.) (Sarstedt, Henseler, & Ringle, 2011). Therefore, considering the overall sample, the FIMIX procedure applied (maximum iterations: 5000, stop criterion: 10) to explore unobserved heterogeneity of the data. After that, the PLS Algorithm procedure (maximum iterations: 300, stop criterion: 7, missing values: mean replacement) was performed to determine the significance levels of the loadings, weights and path coefficients. Bootstrapping procedure (subsamples: 5000) was also applied to determine the significance levels of the proposed hypotheses. Finally, the Q^2 blindfolding procedure was used to identify and assess the accuracy of the tested hypotheses.

Apart from the statistical analysis of the overall sample, it was also applied a multi-group analysis (MGA) to compare the subsamples. The MGA was running two times; the participants of Nijmegen University are compared with the participants of Utrecht University, and the participants from Languages and Communication field are compared with participants from Other Disciplines. MGA indicates if the model is the same or different between the two groups (Garson, 2016). Previous research conducted MGA intends either to pool the data (Henseler, Ringle, & Sarstedt, 2016), or to ensure that the moderating effects and not the variance of the model are responsible for the possible differences in path coefficients (Teo, Lee, Chai, & Wong, 2009). In this study, MGA was conducted firstly to investigate if the model is the same between the groups and thus possible differences in their structural paths to be attributed to moderating effects, and secondly to confirm the possibility of pooling the sample, which also is supported by the unobserved heterogeneity test. Before running MGA, the establishment of measurement invariance is essential, as it indicates if the inner model constructs measure the same thing (Garson, 2016: 168). As stated by Hair, Sarstedt, Ringle, & Gudergan (2018), group comparisons can be misleading unless researchers establish the invariance of their measures, as the differences between groups may be attributed to the variance of the measurement model and not to the differences between the groups. Therefore, in smartPLS, the permutation algorithm was running to be employed the measurement invariance (MICOM) test, and then the multi-group analysis was performed.

6. Data Analysis and Results

In this section, the analysis of the data and the results of the research are discussed. Firstly, the data for the overall sample is presented, followed by the multi-group analysis. The chapter starts with the assessment of unobserved heterogeneity and the alterations that were made for the increase of reliability and validity of the model. After that, the descriptive statistics of constructs are presented, and the measurement and structural model for the overall sample are described. Finally, the results of the multi-group analysis are presented.

6.1 Unobserved Heterogeneity

Before starting the assessment of the research model, the examination of the heterogeneity of the sample was necessary. As the overall sample came from the pooling of different groups, it had to be tested for unobserved heterogeneity.

The sample was tested for unobserved heterogeneity by running FIMIX-PLS. With this procedure, researchers calculate likelihood-based information criteria, which provide an indication of how many segments to retain from the data. FIMIX-PLS algorithm was used on the data 10 times each, using consecutive numbers of groups *s* and starting with the one-segment solution (Sarstedt & Ringle, 2010). The range of possible segment numbers is related to the interplay between the sample size and the minimum sample size requirements (30 participants according to the rule of thumb) to reliably estimate the given model (Hair et al., 2014). In our model, the maximum number of segments could be 2, as our sample consists of 85 participants.

Following Sarstedt et al. (2011), we based our selection of the adequate number of subgroups on several criteria. Specifically, we relied on the Akaike information criterion (AIC), modified AIC3, Bayesian information criterion (BIC), and heuristic consistent AIC (CAIC) (Sarstedt & Ringle 2010) (Table 2). The smaller the value of a certain information criterion, the better the segmentation solution. The information criteria and the segment size pointed 1 segment as a preferable solution. BIC and CAIC have a smaller value for one segment instead of two segments (AIC3, BIC, and CAIC). Furthermore, with two subgroups we derived a larger group with $\pi1$= 0.832 and a smaller group with $\pi2$= 0.168 (Table 2), which means that the second group is under the minimum sample size requirements (30 participants). Hair, Sarstedt, Matthews, & Ringle (2016: 70) support that if the analysis yields an extraneous segment that is too small to warrant valid analysis, the researcher should decrease the number of segments. Thus, the FIMIX-PLS Algorithm points out that our sample is not necessary to be divided into segments, which means that unobserved heterogeneity is not a critical issue for our model.

Table 2 FIMIX-PLS evaluation criteria and relative segment sizes πG.

S	Akaike information criterion (AIC)	Modified AIC3	Bayesian information criterion (BIC)	Consistent AIC (CAIC)	Relative segment sizes πG	
					g=2	g=2
S=1	174.199	181.199	191.215	198.215	1.000	
S=2	159.350	174.350	195.813	210.813	0.832	0.168

6.2. First and revised research model

Before the analysis of the final version of the research model, some necessary alterations were performed for the increase of reliability and validity of the model. Testing the reliability of indicators, the outer loadings of facilitating conditions' indicator FC Q_7.3 (3.041) and indicator FC Q_6.1 (0.324) were lower that the threshold value of 0.40 (Appendix 1). Therefore, they should be eliminated from the construct (Hair et al., 2014). Furthermore, examining inner variance inflation factor (VIF) values for collinearity issues on constructs, the value of At (6.087) was greater than the threshold value of 5. (Henseler et al., 2016). At presented multicollinearity with HM (4.079) and PE (3.014), as the values of HM and PE were high, compared to the other constructs' values (Table 3, First Model). Hair et al. (2014: 205) recommend the elimination of construct, when it indicates a critical level of collinearity. To decide which construct should be removed, we tested inner VIF values multiple times, dropping each time one of the problematic constructs. Finally, the construct of HM was selected for removing, as the variance inflation factor (VIF) values of all the remaining constructs remained below 5 (Table 3). Therefore, the following analysis is based on the revised model, which does not include the two indicators of FC and the construct of HM (Figure 2).

Table 3 Inner VIF values first and revised model.

	First Model	**Revised Model**
	BI	**BI**
BI	-	-
HM	4.079	-
At	6.087	3.715
FC	1.559	1.586
SI	1.382	1.419
Ht	1.598	1.457
EE	2.987	2.654
PE	3.014	3.139

BI behavioral intention, HM hedonic Motivation, At attitude, FC facilitating conditions, SI social influence, Ht habit, EE effort expectancy, PE performance expectancy.

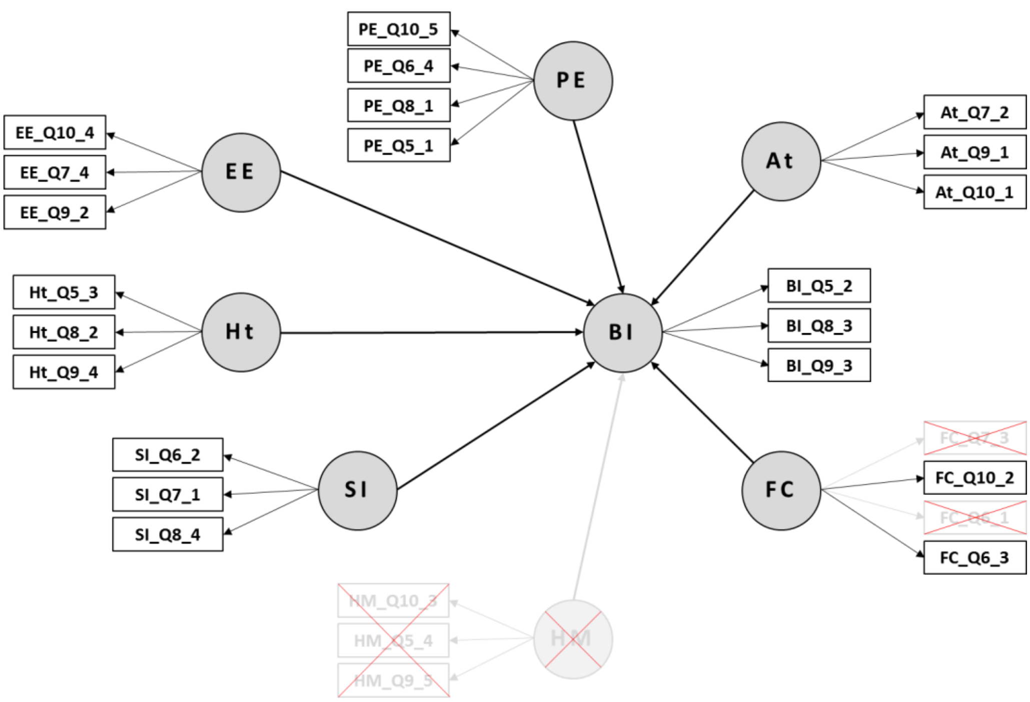

Fig. 2 The revised and final research model.

6.3. Descriptive analysis

In this section, the descriptive analysis of constructs is presented considering both the overall as well as the individual groups' sample.

The descriptive analysis of constructs for the overall sample is shown in Table 4. The means of BI, SI, and Ht are close to the midpoint of 4.00 (the lowest 3.29), which represents a neutral perspective, while the means of At, FC, EE and PE are above the midpoint, ranging from a low of 4.93 to a high of 6.21. These results indicate that the majority of students generally express positive answers to the variables used in the research model. Especially, FC (5.81) and EE (6.21) are the most positively rated variables, as indicated by their high means. The standard deviations ranged from .78 to 1.56 indicating a narrow spread around the mean.

Because of the maximum likelihood estimation used in this study in structural equation modeling (SEM), it is crucial to test the assumption of normality of variables (Curran, West, & Finch, 1995). Following the assumption of normality (i.e., skewness < 3; kurtosis < 10) proposed by Kline (2005), the variables could be regarded as fairly normal for further analyses, as the skew index ranges from -1.21 to 0.25 and kurtosis index ranges from -1.20 to 2.46.

Following to group samples, Table 5 shows the descriptive statistics of constructs (mean and standard deviation) and the results from the two-tailed independent t-test that compared the mean value of constructs as rated by the participants from Nijmegen and the participants from Utrecht University. The results indicate that the group from Nijmegen University responded more positively to the variables compared to the group from Utrecht University. Explicitly, the mean scores of BI, At, Ht, EE and PE are statistically significantly different between the two groups ($p<0.05$); Nijmegen's group presents greater mean values to these constructs comparing to Utrecht's group. Furthermore, the mean value of PE presents the largest difference between the two groups (mean= 5.50 for Nijmegen's group and mean = 4.39 for Utrecht's group). On the other hand, there was no statistically significant difference between the two groups for the mean scores of FC and SI.

In addition, Table 6 shows the descriptive statistics of constructs for the group of Languages and Communication and the group of Other Disciplines, as well as the results from the t-test that compared the mean values of constructs rated by the two groups. All mean scores of constructs do not present a statistically significant difference indicating that both groups rated the constructs in the same way.

Table 4 Descriptive Statistics of the constructs for the <u>overall sample</u>

Constructs	Mean	Std. Deviation	Skewness	Kurtosis
BI	3.71	1.56	-0.01	-1.20
At	5.12	1.09	-1.15	2.46
FC	6.21	0.78	-1.20	1.07
SI	3.87	1.18	-0.38	-0.49
Ht	3.29	1.51	0.25	-1.07
EE	5.81	0.89	-1.21	2.13
PE	4.93	1.18	-0.95	0.53

BI behavioral intention, At attitude, FC facilitating conditions, SI social influence, Ht habit, EE effort expectancy, PE performance expectancy

Table 5 Descriptive Statistics of the constructs and t-test for Nijmegen – Utrecht University.

	Mean		**Std. Deviation**		**t-test**	
	Nijmegen University	**Utrecht University**	**Nijmegen University**	**Utrecht University**	**t**	**P-value**
BI	4.02	3.25	1.55	1.52	2.30*	0.024
At	5.52	4.61	0.92	1.20	4.22*	0.000
FC	5.29	5.14	0.77	0.72	0.91	0.365
SI	3.73	4.08	1.22	1.25	-1.36	0.178
Ht	3.73	2.75	1.54	1.44	2.72*	0.008
EE	6.06	5.44	0.76	1.11	3.13*	0.003
PE	5.50	4.39	1.01	1.20	4.72*	0.000

Nijmegen, n= 48, Utrecht, n= 36
***p < 0.05**
BI behavioral intention, At attitude, FC facilitating conditions, SI social influence, Ht habit, EE effort expectancy, PE performance expectancy

Table 6 Descriptive Statistics and t-test for Languages and Communication – Other Disciplines.

	Mean		**Std. Deviation**		**t-test**	
	Languages and Communication	**Other Disciplines**	**Languages and Communication**	**Other Disciplines**	**t**	**P-value**
BI	3.57	3.73	1.48	1.61	-0.46	0.650
At	4.95	5.24	0.94	1.21	-1.16	0.248
FC	5.06	5.12	0.64	0.70	-0.38	0.704
SI	4.00	3.72	1.03	1.27	1.08	0.282
Ht	3.20	3.30	1.55	1.50	-0.30	0.767
EE	5.69	5.87	0.83	0.93	-0.93	0.353
PE	4.74	5.07	1.18	1.20	-1.24	0.218

Languages and Communication, n= 36, Other Disciplines, n= 46
***p < 0.05**
BI behavioral intention, At attitude, FC facilitating conditions, SI social influence, Ht habit, EE effort expectancy, PE performance expectancy

6.4. Measurement Model

The assessment of measurement model relies on the establishment of convergent validity and discriminant validity. To begin with the convergent validity, it is necessary to assess the reliability of questions, the composite reliability of constructs and the average variance extracted (AVE) by constructs (Fornell & Larcker, 1981). Reliability of the items for the complete model was tested by examining the outer loadings of each question, as the model is reflective. All indicators have an adequate value considering the outer loadings, as it is either in the acceptable range 0.40-0.70 or greater than the threshold value of 0.70 (Hair et al., 2014; Hulland, 1999) (Table 7 – Outer Loadings).

Furthermore, the internal consistency reliability of the constructs was investigated examining the Composite Reliability and Cronbach's Alpha. Most of the constructs had an adequate score of Cronbach's Alpha and Composite Reliability, as it was greater than 0.7 (Hair et al., 2014). Only the Composite reliability and Cronbach's Alfa of FC present problematic values (0.825 and 0.591, respectively), but it is not a critical issue. Hair et al. (2014: 137) supports that the true internal consistency reliability usually lies between Cronbach's Alpha (representing the lower bound as a conservative measure) and the composite reliability (representing the upper bound) (Table 7 - Cronbach's Alpha, Composite reliability). Therefore, it is considered that the internal consistency reliability is established in our model.

The final indicator of convergent validity is the average variance extracted (AVE), which tests the amount of variance captured by the construct in relation to the amount of variance attributable to measurement error (Fornell & Larcker, 1981). After dropping the two indicators of FC (see above), the AVE value of all constructs exceed the threshold of 0.50 (Table 7 - AVE), and thus the convergent validity of constructs is adequate.

Finally, the discriminant validity was tested to explore the extent that the constructs differ. Discriminant validity was assessed for the complete sample by looking at heterotrait-monotrait ratio (HTMT) of the correlations, as proposed by Henseler et al. (2016). However, as PLS-SEM is not based on any distributional assumptions, standard parametric significance tests cannot be applied to test whether the HTMT values are significantly different from 1. Thus, Hair et al. (2014: 141) recommend the computing bootstrap confidence intervals. After running bootstrapping, confidence interval results of the HTMT criterion ensure the discriminant validity of constructs, as neither of the confidence intervals includes the value 1 (Table 7 - HTMT Confidence Intervals).

Table 7 Evaluation of Measurement model.

Constructs	Outer Loadings	Cronbach's Alpha	Composite reliability	AVE
BI				
BIQ52	0.921			
BIQ83	0.962	0.950	0.968	0.910
BIQ93	0.956			
At				
ATQ101	0.919			
ATQ72	0.866	0.881	0.926	0.807
ATQ91	0.892			
FC				
FCQ102	0.747	0.591	0.825	0.704
FCQ63	0.610			
SI				
SIQ62	0.664			
SIQ71	0.868	0.716	0.832	0.626
SIQ84	0.830			
HT				
HtQ53	0.846			
HtQ82	0.826	0.823	0.894	0.739
HtQ94	0.886			
EE				
EEQ104	0.910			
EEQ74	0.561	0.761	0.858	0.683
EEQ92	0.961			
PE				
PEQ105	0.824			
PEQ51	0.725	0.837	0.891	0.674
PEQ64	0.853			
PEQ81	0.881			

BI behavioral intention, At attitude, FC facilitating conditions, SI social influence, Ht habit, EE effort expectancy, PE performance expectancy

HTMT Confidence Intervals	**5.0%**	**95.0%**
BI -> AT	0.579	0.805
FC -> BI	0.088	0.444
FC -> AT	0.287	0.697
FC -> EE	0.635	1.047
SI -> BI	0.585	0.83
SI -> AT	0.305	0.62
SI -> FC	0.13	0.461
SI -> HT	0.277	0.68
SI -> EE	0.23	0.552
SI -> PE	0.44	0.753
HT -> BI	0.301	0.654
HT -> AT	0.346	0.675
HT -> FC	0.165	0.467
HT -> EE	0.192	0.48
EE -> BI	0.211	0.545
EE -> AT	0.615	0.918
PE -> BI	0.548	0.805
PE -> AT	0.844	0.981
PE -> FC	0.224	0.614
PE -> HT	0.445	0.762
PE -> EE	0.418	0.822

6.5. Structural Model and Hypotheses testing

Satisfactory outcomes for the measurement model are a prerequisite for evaluating the relationships in the structural model (Hair et al., 2014). Following the structural model assessment procedure, we first need to check the structural model for collinearity issues by examining the VIF values of all sets of predictor constructs. As already mentioned, the construct of HM was deleted (6.2.), and thus, all constructs had VIF values below the threshold of 5. (Table 3).

After testing collinearity issues, estimates are obtained for the structural model relationships for the pooled sample (i.e., the path coefficients), which represent the hypothesized relationships among the constructs. As it is shown in Table 8, the relationships between At-BI and SI-BI are statistically significant ($p < 0.05$), and thus, H1 and H3 are supported. At (0.520, $p < 0.05$) and SI (0.439, $p < 0.05$) were found to have a significant positive impact on BI toward using MyST, with At having the strongest magnitude on the relationship with BI. EE (- 0201, $p < 0.05$) had a significant but negative effect on BI, which means that H5 is not supported. Furthermore, the H2, H4 and H6 are not supported, as the results showed that the impact of FC, Ht and PE on BI is not statistically significant ($p > 0.05$).

Considering the in-sample predictive power of our model, it is moderate to substantial, considering the rule of Thumb (Hair et al., 2014), as the coefficient of determination value (R^2) is 0.606 for BI (Table 8). In other words, the basic structure of UTAUT could explain 60.6 % of the variation in students' behavioral intention concerning the pooled sample. In our model, the R^2 value of BI is considered acceptable, as Henseler et al. (2009) support that if the endogenous latent variable relies on several exogenous latent variables, the value should exhibit at least a substantial level. Furthermore, concerning effect size (f^2), no one of the constructs presents a statistically significant effect on BI ($p > 0.05$).

Concerning the out-of-sample predictive power (Q^2) estimated by the blindfolding procedure, the research model presents a high degree of predictive relevance with regard to the BI ($Q^2 = 0.52$), as it is greater than 0.35 (Cohen, 1988). Besides, the effect size q^2, which is a measurement for the relative impact of predictive relevance (Hair et al., 2014: 215), indicates that SI has a medium to a large effect ($q^2 = 0.25$) on BI. Furthermore, At has a medium effect ($q^2 = 0.14$) on BI, while EE has a small effect size ($q^2 = 0.04$). Finally, FC, Ht and PE have no effect on BI, as their values of effect size are 0.00 (Table 8).

Table 8 Analysis of Structural Model

	Path Coefficient	p-value	R^2	f^2	p-value f^2	Q^2	q^2
BI	-	-		-	-	-	-
At -> BI	0.520*	0.001	0.606	0.184	0.083	0.52	0.14
FC -> BI	0.077	0.210		0.038	0.202		0.04
SI -> BI	0.439*	0.000		0.002	0.471		0
Ht -> BI	0.036	0.353		0.010	0.394		0
EE -> BI	-0.201	0.039		-	-		-
PE -> BI	0.051	0.356		0.002	0.469		0

***p<0.05**
BI behavioral intention, At attitude, FC facilitating conditions, SI social influence, Ht habit, EE effort expectancy, PE performance expectancy

6.6. Multi-group analysis (MGA)

After the evaluation of the measurement and structural model for the overall sample, measurement invariance and structural invariance tests were performed to investigate if the research model is the same for the compared groups. Firstly, a measurement invariance (MICOM) test was running to explore if the inner model constructs measure the same thing (Garson, 2016: 168). Secondly, a multi-group analysis (MGA) was conducted to ensure the structural invariance of the model and ensure that the structural model gives the same results for the compared groups. The establishment of measurement and structural invariance is essential, as it allows researchers to confirm that the differences between the groups do not arouse by the research model but by the different characteristics of groups.

6.6.1 Measurement Invariance

The measurement invariance test was performed, focusing on configural invariance, compositional invariance and scalar invariance. Firstly, configural invariance was established between the Nijmegen and Utrecht's group, as the basic factor structure exists in both groups, in terms of a number of constructs and items associated with each construct (Garson, 2016: 186). Furthermore, configural invariance was established between the groups from different faculties, as the same research model was applied in both groups. The compositional invariance was proved between the groups of both categories, as "permutation p-value", which tests if item loadings in the outer model are invariant across groups, are non-significant ($p > 0.05$) (Table 9). As configural and compositional invariance were investigated, partial measurement invariance is established. Following the recommendation of Garson (2016: 186), when partial measurement invariance is established, the investigation of scalar invariance is not necessary for purposes of comparing models. Therefore, we proceed directly with MGA to compare structural paths across groups.

Table 9. Configural Invariance.

	Nijmegen-Utrecht		Ling./Commun.-Other Disciplines	
	Original Correlation	**Permutation p-Values/Step 2**	**Original Correlation**	**Permutation p-Values**
BI	1,000	0,298	1,000	0,315
At	1,000	0,734	0,998	0,233
FC	0,993	0,861	0,906	0,585
SI	0,993	0,448	0,977	0,137
Ht	0,984	0,236	0,998	0,774
EE	0,998	0,931	0,976	0,610
PE	0,996	0,506	0,999	0,842

***p< 0.05**
BI behavioral intention, At attitude, FC facilitating conditions, SI social influence, Ht habit, EE effort expectancy, PE performance expectancy

6.6.2 PLS-MGA

A multi-group analysis was performed to examine the differences in the structural model between the groups of participants (Nijmegen – Utrecht and Languages and Communication – Other Disciplines). Table 10 shows only the results from the comparison of path coefficients of two groups, as the comparison of descriptive statistics of constructs was implemented in the section of descriptive analysis (6.3, Table 5 and 6). The results show that there is no statistically significant difference in the structural paths across groups, as the p-values are greater than 0.05 (Table 10). In other words, possible differences in path coefficients between the two groups can be attributed to moderating effects and not to the structural variance of the model, as the structural model provides the same results for both groups (Hair et al., 2018).

Table 10. PLS - MGA

	Path Coefficients-diff Nijmegen - Utrecht	p-Value (Nijmegen vs Utrecht)	Path Coefficients-diff Ling/Comm.-Other Disciplines	p-Value Ling./Comm. Vs Other Disciplines
BI	-	-	-	-
At -> BI	0,356	0,866	0,039	0,547
FC -> BI	0,051	0,605	0,300	0,079
SI -> BI	0,086	0,324	0,243	0,087
Ht -> BI	0,063	0,625	0,321	0,944
EE -> BI	0,270	0,154	0,009	0,489
PE -> BI	0,166	0,256	0,034	0,463

***p < 0.05** BI behavioral intention, At attitude, FC facilitating conditions, SI social influence, Ht habit, EE effort expectancy, PE performance expectancy

After ensuring the measurement and structural invariance of the model for the compared groups, we presented the relationships of constructs with BI for the compared groups (Table 11). Both the group from Nijmegen and Utrecht University show a positive and statistically significant relationship between the constructs of SI and BI ($p < 0.05$). On the other hand, only the group from Utrecht indicates a positive and statistically significant relationship of At with BI (0,673, p-value < 0.05). Furthermore, the rest of the constructs, namely FC, Ht, EE and PE, do not have a statistically significant effect on BI of both groups. Finally, the basic structure of UTAUT2 could explain 59.2 % (R^2 = 0,592) of the variation in students' behavioral intention at Nijmegen group, while it explains 62,7 % (R = 0.627) of the variation at the Utrecht group.

Considering the group of Languages and Communication with the group of Other Disciplines, it is observed that the relationship between At-BI and SI-BI is statistically significant for both groups ($p < 0.05$) (Table 11). However, at the group of Languages and Communication, At (0.561) and SI (0.522) have a similar effect on BI, while at the group of Other Disciplines, the At (0.600) is a stronger indicator of BI comparing to SI (0.279). The rest of the constructs, namely FC, Ht, EE and PE, do not have a statistically significant effect on BI of both groups. Finally, variance explained was higher for Languages and Communication group (R^2 = 0.717, 71.7%) comparing to the variance of Other Disciplines group (R^2 = 0.616, 61.6%).

Table 11 Structural Paths of groups and R^2.

	Nijmegen (n=48)		**Utrecht (n=36)**		**Languages and Communication (n=36)**		**Other Disciplines (n=46)**	
	Path Coefficient	P values	Path Coefficient	P values	Path Coefficient	P values	Path Coefficient	P values
BI	-	-	-	-	-	-	-	-
At -> BI	0.317	0.057	0.673*	0.004	0.561*	0.003	0.600*	0.010
FC -> BI	0.063	0.320	0.115	0.228	0.179	0.068	-0.143	0.227
SI -> BI	0.532*	0.000	0.447*	0.000	0.522*	0.000	0.279*	0.028
Ht -> BI	-0.011	0.466	0.052	0.375	-0.125	0.204	0.196	0.078
EE -> BI	-0.040	0.374	-0.311	0.098	-0.162	0.215	-0.17	0.206
PE -> BI	0.075	0.330	-0.091	0.332	0.076	0.330	0.042	0.435
R^2	0.592		0.627		0.717		0.616	

***p < 0.05**
BI behavioral intention, At attitude, FC facilitating conditions, SI social influence, Ht habit, EE effort expectancy, PE performance expectancy

7. Discussion

Due to the advancements on technological field, more and more CALL systems have been developed and used in academic contexts. However, ASR-based CALL systems are still under development, and they are not often used in academic settings. The research reported on in this paper addressed the evaluation of a CALL system, called *My Speech Trainer* that employs ASR technology and offers practice and feedback on academic English oral skills. The novelty of the system relies on the combination of ASR technology with the context of academic language. We managed to develop such a system, to adapt it to different academic courses, and to test it in realistic conditions.

A significant gap in the literature is that most of the studies focus on ASR performance neglecting the study of user's acceptance and the use of the system in realistic conditions. Furthermore, little research exists for the investigation of user's intention to use ASR-based systems by testing and applying the revised version of UTAUT2 model in academic settings. Along that line, this research investigating the student's acceptance of MyST contributes to the existing literature in the field of ASR - CALL as well as in the field of technology acceptance model. Specifically, the research model UTAUT2 extended by the construct of attitude (At) was developed to assess the behavioral factors that influence user's intention to use the MyST, as well as organizational and individual factors that affect the relation between behavioral constructs and intention to use the tool.

The results of data analysis showed that SI turned out to be the most significant factor of intention to use the system both for the overall sample as well as for all the individual groups. This is an interesting finding, as it indicates that SI appears to be an important determinant of BI and independent from the settings of tool's use. In contradiction with the study of Sun and Zhang (2006) that supports that SI has a significant influence on BI in mandatory, but not in voluntary settings, in our study, the moderating effect of organizational factors did not affect the relationship of SI with BI. Although the users from Utrecht used the system in a semi-voluntary context and more systematically than the users of Nijmegen that used the tool voluntarily, the SI remains a strong predictor for both groups. These findings are in accordance with many previous studies in acceptance of technology for language learning in spite of different characteristics of participants and different contexts of research (Hsu, 2012; Khechine, 2014; Tan, 2013). Furthermore, concerning the faculty groups, At (0.561) and SI (0.522) have almost the same effect on BI of Lanaguages and Communication group, as their path coefficients have similar values. On the other hand, at the group of Other Disciplines, At (0.600) is the strongest predictor of BI, and it presents a greater value of path coefficient comparing to SI (0.279). The content of the tool may be related to this difference; users studied Languages and Communication may consider that the device could help not only to academic English but also to their studies, and thus they had a greater intention to be conformed with the important other's opinion and especially with their professor's recommendations. Considering the value of these results in CALL systems' acceptance, it seems that the acceptance could be increased by increasing SI and developing specific course content. Therefore, CALL tools should be combined with standard courses at university and be recommended by teachers or other important persons.

This study further found that At is an important predictor of BI, as it presents a positive and statistically significant relationship with BI both of the pooled sample and of all groups, with the only exception the group of Nijmegen. Considering the moderating effect of organizational factors, participants from Nijmegen performed activities with general content, and they did not use the MyST in the framework of an academic course. On the other hand, participants from Utrecht University used the MyST in the framework of a Linguistic course, and they did activities with specific course content. Therefore, the students from Utrecht may intend to use the MyST in the future, because they realized that using MyST during their academic course makes the learning of course content easier and more pleasant. On the other hand, although the students from Nijmegen showed a positive attitude toward using the MyST, as indicated by the high mean value of the construct (5.52), At is not a predictor of behavioral intention. This difference compared to the Utrecht's group may be attributed to the independence of participants from Nijmegen from an academic course context.

Furthermore, considering the results for PE, it does not present a statistically significant relationship either for the overall sample or for the subsets. If the users do not use the MyST for a long time, maybe they are not able to understand its usefulness in the development of spoken academic English. The study of Pynoo (2011) indicates that there are interesting changes in the relation of constructs with the BI throughout the use of a tool for a year. During the first month of the use, PE and SI were the predictors of behavioral intention, while after some months of use, PEwas the only significant predictor of BI.

In addition, the results of this study indicate no significant influence of FC on BI to use the MyST for both the overall sample and the subsets. These findings support the hypothesis by Venkatesh et al. (2003), that when both PE and EE are present, FC is not significant in predicting BI of use of technology. However, there are conflicting results for the FC in studies conducted in educational settings. Many studies found a significant relation between FC and BI in a higher educational context (Meng & Wang, 2012; Nair et al., 2015; Tosuntaş, Karadağ, & Orhan, 2015), while others showed opposite results (Teo & Noyes, 2014). It is reported that the willingness of students to use the system depends on the facilitating conditions of systems' infrastructure and of time for training and system's use (Lu, Liu, Yu, & Wang, 2008). Our research shows different results, as the mean value of FC for the group of Nijmegen does not present a statistically significant difference with the mean value of Utrecht's group. We would expect that participants from Nijmegen would have a greater mean value for FC, as they received some instructions for the use of the MyST, and they used it for a long time. The differences in research concerning FC may be due to the different characteristics of participants that used in studies. For example, Venkatesh et al. (2003) conducted research in a business context, while other studies focused on educational settings. Therefore, future research should examine possible differences between business and educational users in the UTAUT model to provide greater insights on the validity of this model.

Apart from the aforementioned, the results of this research are also interesting for the field of technology acceptance model. Different from most of the studies that applied the UTAUT model in ICT research, our study supports theoretically and empirically the UTAUT2 reliability and validity in an academic educational context and more specifically in the CALL and ASR-technology research. This study validates the revised version of the UTAUT model, the UTAUT2 as well as the extension of UTAUT model with At proposed by Dwivedi, Rana, Jeyaraj, Clement & Williams (2017). The present study indicates that At is one of the most significant constructs that directly affects BI to use MyST. Therefore, it supports the findings of latest research (Dwivedi et al., 2017), while it is in contradiction with the previous research, which does not include At in the research model, because the researchers did not observe any significant relationship between At and intention in their statistical analysis (Venkatesh et al., 2003; Venkatesh et al., 2012).

Furthermore, although UTAUT has been validated empirically, research tends to add and investigate the role of different factors and behavioral constructs to overcome the limited explanatory power of such models. This study moves in that direction, and it shows interesting results; adding to the UTAUT2 model the attitude, the model indicates a moderate to substantial explanatory power, as it accounted for 60,6 % of the variance in usage intention. In addition, adding the moderating effects, in most of the cases the variance explained BI was increased, as the proposed model explained 62.7 % of the variance in the BI for the group from Utrecht, 71.7% for the group of Language and Communication and 61.6 % for the group of Other Disciplines. These results indicate that the moderating effects increase the explanatory power of the model, and they are in accordance with other studies (Sun & Zhang, 2006; Venkatesh et al., 2003). Therefore, the previous conceptual model can be adopted in other educational environments, which concentrate on users' (students and lecturers) behavioral intention to use ASR-technologies for academic English. In addition, the high increase of explanatory power that observed at the group of Languages and Communication indicates that the faculty is a strong moderating effect. Further research should be conducted for its influence in user's acceptance of CALL in the academic context, as to the best of our previous knowledge, research investigates the educational background and not the field of studies as moderating factor (Dečman, 2015).

In overall, the positive comments of students that collected from open questions at the end of the UTAUT questionnaire, the high mean value of At and its significant relationship with BI show that MyST not only is accepted by the target users but also they enjoy and they find easy and useful to use it. This is in contradiction with the findings of Choi (2014), who suggested that even the most advance neuro-linguistic programme technology does not adequately examine learners' output and thus it may cause frustration and reduce motivation. In other words, MyST and its ASR-technology may have a high-quality level, and future research could focus on its further development.

8. Limitations and Future Research

Heading to the end of the current study, the limitations faced in our research indicate topics that might be useful to be addressed by future research. First of all, although PLS can deal with small sample sizes (Henseler et al., 2009), the small number of participants may restrict the generalizability of the results, as the p-values are highly dependent on the sample size (Field, 2013). Future research could apply the research model to a complementary, possibly larger sample to expand further the findings of this study. Secondly, our study did not take into account the actual use of MyST, as we did not manage to collect data for the actual use of the system from the Nijmegen group (see Introduction). Therefore, BI is the only measurement that predicts the future use of MyST. Future research could implement the actual use as dependent variable retrieving realistic data about the time of use by the MyST platform.

Furthermore, our study used a cross-sectional design, as it measures the BI of participants at a specific point in time. Given that one's experiences shape perceptions and intentions and change over time, researchers could design a longitudinal study to investigate possible differences in relationships of constructs at different periods of use the MyST. Another interesting avenue for further research might be the investigation of possible cross-cultural differences in the determinant factors that influence students' acceptance of MyST, as MyST addresses to Universities with international students.

Finally, as SI and At are the most important predictors of BI, university lecturers may have a significant role in the acceptance of MyST. They can improve their students' attitude toward using MyST and motivate the implementation of the system in their departments. Therefore, additional research is needed to explore lecturers' perceptions of MyST and to investigate the challenges that lecturers consider that may face adopting the MyST in teaching.

9. Conclusion

This study extended and applied a research model to investigate theoretically and empirically influencing factors on students' acceptance of *My Speech Trainer*, an ASR-based CALL system that offer practice and feedback on academic English oral skills. Results showed that SI and At are the most significant factor of intention to use the MyST both for the overall sample and all groups, and that moderating factors increased a lot the explanatory power of the model. A good strategy for increasing the acceptance of CALL systems is thus to improve SI and increase the user-friendliness of the system, as it can lead to a better attitude toward the system.

One of the novelties of this study is that it focuses on a new ASR-based CALL system, MyST, that provides practice and feedback on oral academic English. Secondly, the acceptance of MyST was investigating considering not only the different characteristics of users but also the different conditions of use of the MyST; some users completed activities with a general content, in a voluntary context and for a short time, while others used MyST to practice with course-specific content in a semi-voluntary and routinely framework.

Furthermore, the various tests carried out in this study ensured that the final UTAUT2 model extended by attitude was valid and reliable. Therefore, it can be adopted in other educational environments that concentrate on users' (students and lecturers) behavioral intention to use ASR-technologies. In addition, most of the studies using the UTAUT2 model included organizational and individual factors as moderating effects. However, our study extended the meaning of organizational factors adding as a type of task the course-content activities, and it presented the faculty as a new moderating factor. Besides, although multi-group analysis is not a common method in previous literature, this study included multi-group analysis to establish that the measurement model would be equivalent for the compared groups and to ensure that the comparisons between the groups would be meaningful.

Concluding, the overall results for the acceptance of MyST is that users were positive toward the intention to use the tool in the future, which enhance the motivation for research in the direction of its further development and improvement.

References

Adair-Hauck, B., Willingham-McLain, L., & Youngs, B.E. (2000). Evaluating the integration of technology and second language learning. *CALICO Journal*, *17*(2), 269-306.

Afshari, M., Ghavifekr, S., Siraj, S., & Jing, D. (2013). Students' attitudes toward computer-assisted language learning. *Procedia-Social and Behavioral Sciences*, *103* (2), 852-859.

Agudo-Peregrina, A.F., Iglesias-Pradas, S., Conde-Gonzalez, M.A., & Hernandez-Garcıa, A. (2014). Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning. *Computers in Human Behavior, 31*, 542-550.

Ajzen, I. (1991). The Theory of Planned Behavior. Organization Behavior and Human Decision Processes. *Academic Press*, *50*(2) 179-211.

AlAwadhi, S., & Morris, A. (2008). The use of the UTAUT Model in the adoption of E-government services in Kuwait. In *Proceedings of* the *41st Annual Hawaii International Conference on System Sciences* (pp. 219-230*)*. Waikoloa, Big Island, HI, USA: IEEE.

Ali, F., Hussain, K., and Ragavan, N.A. (2014). Memorable customer experience: examining the effects of customers experience on memories and loyalty in Malaysian resort hotels. *Procedia – Social and Behavioral Sciences*, *144* (8), 273-279.

Attuquayefio, S., & Addo, H. (2014). Using the UTAUT model to analyze students' ICT adoption. *International Journal of Education and Development using ICT*, *10*(3), 75-86.

Ayres, R. (2002). Learner attitudes towards the use of CALL. *Computer-assisted language learning*, *15*(3), 241-249.

Baltaci-Goktalay, S., & Ozdilek, Z. (2010). Pre-service teachers' perceptions of web 2.0 technologies. *Procedia-Social and Behavioral Sciences*, *2*(2), 4737-4741.

Browne, T., Jenkins, M., & Walker, R. (2006). A longitudinal perspective regarding the use of VLEs by higher education institutions in the United Kingdom. *Interactive Learning Environments*, *14*(2), 177–192.

Cavus, N., & Ibrahim, D. (2009). m-learning: An experiment in using SMS to support learning new English language words. *British Journal of Educational Technology*, *40*(1), 78-91.

Chin, W. W., Marcolin, B. L., & Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic mail emotion/adoption study. *Information systems research*, *14*(2), 189-217.

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.

Choi, I.C. (2016). Efficacy of an ICALL tutoring system and process-oriented corrective feedback. Computer Assisted Language Learning, *29*(2), 334-364.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*(1), 16–29.

Davis, F. D., Bogozzi, R., P., & Warshaw, P., R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, *35* (8), 982-1003.

Dečman, M. (2015). Modeling the acceptance of e-learning in mandatory environments of higher education: The influence of previous education and gender. *Computers in human behavior*, *49*, 272-281.

Dwivedi, Y. K., Rana, N. P., Chen, H., & Williams, M. D. (2011, September). A Meta-analysis of the Unified Theory of Acceptance and Use of Technology (UTAUT). In *IFIP international working conference on governance and sustainability in information systems-managing the transfer and diffusion of it* (pp. 155-170). Springer, Berlin, Heidelberg.

Dwivedi, Y. K., Rana, N. P., Jeyaraj, A., Clement, M., & Williams, M. D. (2017). Re-examining the unified theory of acceptance and use of technology (UTAUT): Towards a revised theoretical model. *Information Systems Frontiers*, *21*(3) 1-16.

Elimat, A. K., & AbuSeileek, A. F. (2014). Automatic speech recognition technology as an effective means for teaching pronunciation. *JALT CALL Journal*, *10*(1), 21-47.

El-Masri, M., & Tarhini, A. (2017). Factors affecting the adoption of e-learning systems in Qatar and USA: Extending the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2). *Educational Technology Research and Development*, *65*(3), 743-763.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. sage.

Foote, J. A., & McDonough, K. (2017). Using shadowing with mobile technology to improve L2 pronunciation. *Journal of Second Language Pronunciation*, *3*(1), 34-56.

Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing Research, 18*(3), 382-388.

Garson, G. D. (2016). Partial least squares: Regression and structural equation models. *Asheboro, NC: Statistical Associates Publishers*.

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.

Hair, Jr, J. F., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2014). *A primer on partial least squares structural equation modeling (PLS-SEM).* Thousand Oaks, CA: Sage Publications.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2013). Partial least squares structural equation modeling: Rigorous applications, better results and higher acceptance. *Long range planning, 46*(1-2), 1-12.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2018). *Advanced Issues in Partial Least Squares Structural Equation Modeling (PLS-SEM)*. Thousand Oaks, CA: Sage.

Hair, Jr, J. F., Sarstedt, M., Matthews, L. M., & Ringle, C. M. (2016). Identifying and treating unobserved heterogeneity with FIMIX-PLS: part I–method. *European Business Review*, *28*(1), 63-76.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2016). Testing measurement invariance of composites using partial least squares. *International marketing review*, *33*(3), 405-431.

Henseler, J., Ringle, C., & Sinkovics, R. (2009). The use of partial least squares path modeling in international marketing. *Advances in International Marketing* (*AIM*)*, 20,* 277-320.

Hong, S, Thong J., and Tam, K. (2006). Understanding continued information technology usage behavior: a comparison of three models in the context of mobile internet. *Decision Support Systems*, *42*(3), 1819-1834.

Hsu, L. (2016). An empirical examination of EFL learners' perceptual learning styles and acceptance of ASR-based computer-assisted pronunciation training. *Computer Assisted Language Learning*, *29*(5), 881-900.

Hulland, J. (1999). Use of partial least squares (PLS) in strategic management research: a review of four recent studies. *Strategic management journal*, *20*(2), 195-204.

Kang, M., Liew, B. Y. T., Lim, H., Jang, J., & Lee, S. (2015). Investigating the determinants of mobile learning acceptance in Korea using UTAUT2. In *Emerging issues in smart learning* (pp. 209-216). Springer, Berlin, Heidelberg.

Khechine, H., Lakhal, S., Pascot, D., & Bytha, A. (2014). UTAUT model for blended learning: The role of gender and age in the intention to use webinars. Interdisciplinary Journal of E-Learning and Learning Objects, *10(*1), 33-52.

Kim, S. S., & Malhotra, N. K. (2005). A longitudinal model of continued IS use: An integrative view of four mechanisms underlying postadoption phenomena. *Management Science, 51*(5), 741-755.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.

Lai, C., Wang, Q., Li, X., & Hu, X. (2016). The influence of individual espoused cultural values on self-directed use of technology for language learning beyond the classroom. *Computers in Human Behavior*, *62*(1), 676-688.

Limayem, M., & Cheung, C. M. (2008). Understanding information systems continuance: The case of Internet-based learning technologies. *Information & Management, 45*(4), 227-232.

Limayem, M., Hirt, S. G., & Cheung, C. M. (2007). How habit limits the predictive power of intention: The case of information systems continuance. *MIS Quarterly*, *31*(4), 705-735.

Liu, Y. C., & Huang, Y.-M. (2015). Using the UTAUT Model to Examine the Acceptance Behavior of Synchronous Collaboration to Support Peer Translation. *JALT CALL Journal*, *11*(1), 77–91.

Lopez-Nicolas, C., Molina-Castillo, F. J., & Bouwman, H. (2008). An assessment of advanced mobile services acceptance: Contributions from tam and diffusion theory models. *Information & Management*, *45*(6), 359–364.

Lowenthal, J. (2010). Using mobile learning: determinates impacting behavioral intention. *The American Journal of Distance Education*, *24*(4), 195-206.

Lu, J., Liu, C., Yu, C. S., & Wang, K. (2008). Determinants of accepting wireless mobile data services in China. *Information & Management*, *45*(1), 52-64.

Mah, B.Y., & Er, A.N. (2009). Writing web logs in the ESL classroom: A study of student perceptions and the Technology Acceptance Model. *Asian Journal of University Education* (*ACRULeT*), *5*(1), 47-70. Retrieved from http://eprints.uitm.edu.my/id/eprint/363.

McCrocklin, S. (2019). ASR-based dictation practice for second language pronunciation improvement. *Journal of Second Language Pronunciation*, *5*(1), 98-118.

Meng, H. J., & Wang, T. Y. (2018). Acceptance of IWBs instruction and concomitant behavior through self-regulation learning. *GSTF Journal on Computing (JoC)*, *1*(4), 210-215.

Morosan, C., & DeFranco, A. (2016). It's about time: revisiting UTAUT2 to examine consumers' intentions to use NFC mobile payments in hotels. *International Journal of Hospitality Management*, *53*(1), 17–29.

Nair, P. K., Ali, F., & Leong, L. C. (2015). Factors affecting acceptance & use of ReWIND: Validating the extended unified theory of acceptance and use of technology. *Interactive Technology and Smart Education*, *12*(3), 183-201.

Nassuora, A. B. (2012). Students acceptance of mobile learning for higher education in Saudi Arabia. *American Academic & Scholarly Research Journal*, *4*(2), 24-30.

Neri, A., Cucchiarini, C., & Strik, H. (2001). Effective feedback on L2 pronunciation in ASR-based CALL. In *Proceedings of the Workshop on Computer Assisted Language Learning, Artificial Intelligence in Education Conference* (pp. 40-48). San Antonio, TX: AIED.

Neri, A., Cucchiarini, C., & Strik, H. (2003). Automatic speech recognition for second language learning: how and why it actually works. In *Proceedings of International Congress of Phonetic Sciences* (pp. 1157-1160). Barcelona, IPA.

Neri, A., Cucchiarini, C., & Strik, H. (2006). ASR-based corrective feedback on pronunciation: Does it really work? P*roceedings of International Conference on Spoken Language Processing* (pp. 1982–1985). Pittsburgh, PA.

Neri, A., Mich, O., Gerosa, M., & Giuliani, D. (2008). The effectiveness of computer-assisted pronunciation training for foreign language learning by children. *Computer Assisted Language Learning*, *21*(5), 393-408.

Padilla-Meléndez, A., Aguila-Obra, A. R. D., & Garrido-Moreno, A. (2013). Perceived Playfulness, Gender Differences and Technology Acceptance Model in a Blended Learning Scenario. *Computers & Education*, *63* (1), 306-317.

Sarstedt, M., Henseler, J., & Ringle, C. M. (2011). Multigroup analysis in partial least squares (PLS) path modeling: Alternative methods and empirical results. In *Measurement and research methods in international marketing*. Emerald Group Publishing Limited *Technology Studies*, *2* (2), 283-309.

Savvani, S. (2018). State-of-the-Art Duolingo Features and Applications. In *International Conference on Interactive Collaborative Learning* (pp. 139-148). Springer, Cham.

Schepers, J., & Wetzels, M. (2007). A meta-analysis of the technology acceptance model: Investigating subjective norm and moderation effects. *Information & Management, 44*(1), 90–103.

Soleimani, E., Ismail, K., & Mustaffa, R. (2014). The acceptance of mobile assisted language learning (MALL) among postgraduate ESL students in UKM. *Procedia-Social and Behavioral Sciences*, *118*, 457-462.

Sun, H., & Zhang, P. (2006). The role of moderating factors in user technology acceptance. *International journal of human-computer studies*, *64*(2), 53-78.

Tan, P. J. B. (2013). Applying the UTAUT to understand factors affecting the use of English e-learning websites in Taiwan. *Sage Open*, *3*(4), 1-12.

Tarhini, A., Hone, K., & Liu, X. (2013). User acceptance towards web-based learning systems: Investigating the role of social, organizational and individual factors in European higher education. *Procedia Computer Science*, *17* (1), 189-197.

Tarhini, A., Hone, K., Liu, X., & Tarhini, T. (2016). Examining the moderating effect of individual-level cultural values on users' acceptance of E-learning in developing countries: a structural equation modeling of an extended technology acceptance model. *Interactive Learning Environments*, *25*(3), 306-328. doi:10.1080/10494820.2015.1122635.

Teo, T. (2009). Modelling technology acceptance in education: A study of pre-service teachers. *Computers & Education*, *52*(2), 302-312.

Teo, T., & Noyes, J. (2014). Explaining the intention to use technology among pre-service teachers: a multi-group analysis of the Unified Theory of Acceptance and Use of Technology. *Interactive Learning Environments*, *22*(1), 51-66.

Teo, T., Lee, C. B., & Chai, C. S. (2008). Understanding pre-service teachers' computer attitudes: applying and extending the technology acceptance model. *Journal of computer assisted learning*, *24*(2), 128-143.

Teo, T., Lee, C. B., Chai, C. S., & Wong, S. L. (2009). Assessing the intention to use technology among pre-service teachers in Singapore and Malaysia: A multigroup invariance analysis of the Technology Acceptance Model (TAM). *Computers & Education*, *53*(3), 1000-1009.

Tosuntaş, Ş. B., Karadağ, E., & Orhan, S. (2015). The factors affecting acceptance and use of interactive whiteboard within the scope of FATIH project: A structural equation model based on the Unified Theory of acceptance and use of technology. *Computers & Education*, *81*, 169-178.

Venkatesh, V. & Davis, F.D., 2000. A theoretical extension of the technology acceptance model: four longitudinal field studies. *Management Science*, *46* (2), 186–204.

Venkatesh, V., & Zhang, X. (2010). Unified theory of acceptance and use of technology: US vs. China. *Journal of global information technology management*, *13*(1), 5–27.

Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Quarterly*, *36*(1), 157-178.

Wang, Y., Wu, M., & Wang, H. (2009). Investigating the determinants and age and gender differences in the acceptance of mobile learning. *British Journal of Educational Technology*, *40(1)*, 92-118.

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56-75). Thousand Oaks, CA, US: Sage Publications

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.

Williams, M. D., Rana, N. P., & Dwivedi, Y. K. (2015). The unified theory of acceptance and use of technology (UTAUT): a literature review. *Journal of Enterprise Information Management*, *28*(3), 443–488. https://doi.org/10.1108/JEIM-09-2014-0088

Yang, S. (2013). Understanding undergraduate students' adoption of mobile learning model: A perspective of the extended UTAUT2. *Journal of convergence information technology*, *8*(10), 969.

Zhao, Y. (2003). Recent developments in technology and language learning: A literature review and meta-analysis. *CALICO Journal*, *21*(1), 7-27.

Appendix 1. Outer Loadings

	AT	BI	EE	FC	HM	HT	PE	SI
ATQ101	0,919							
ATQ72	0,866							
ATQ91	0,892							
BIQ52		0,921						
BIQ83		0,962						
BIQ93		0,956						
EEQ104			0,910					
EEQ74			0,561					
EEQ92			0,961					
FCQ102				0,747				
FCQ61				0,341				
FCQ63				0,610				
FCQ73				0,324				
HMQ103					0,952			
HMQ54					0,902			
HMQ95					0,952			
HTQ53						0,846		
HTQ82						0,826		
HTQ94						0,886		
PEQ105							0,824	
PEQ51							0,725	
PEQ64							0,853	
PEQ81							0,881	
SIQ62								0,664
SIQ71								0,868
SIQ84								0,830

Appendix 2. The UTAUT Questionnaire

no	**Statement**	**Construct**	**Position Qualtrics**
1	BI1. I intend to use MyST in the next 3 months.	Behavioral intention to use the system	Q9_3
2	BI2. I predict I would use MyST in the next 3 months.	Behavioral intention to use the system	Q5_2
3	BI3. I plan to use MyST in the next 3 months.	Behavioral intention to use the system	Q8_3
4	PE1. I would find MyST useful in my studies.	Performance Expectancy	Q5_1
5	PE2. Using MyST would enable me to speak academic English better.	Performance Expectancy	Q10_5
6	PE3. Using MyST would improve my academic English speaking skills.	Performance Expectancy	Q6_4
7	PE4. If I practice English with MyST, I will increase my chances of studying successfully	Performance Expectancy	Q8_1
8	EE1. It would be clear and understandable to me how to use MyST.	Effort Expectancy	Q7_4
9	EE2. I would find MyST easy to use.	Effort Expectancy	Q9_2
10	EE3. Learning to use MyST is easy for me.	Effort Expectancy	Q10_4
11	FC1.A specific person (or a group) are available for assistance with MyST if I need it	Facilitating Conditions	Q6_1
12	FC2. I have the knowledge necessary to use MyST.	Facilitating Conditions	Q10_2
13	FC3. MyST is not compatible with other systems I use.	Facilitating Conditions	Q7_3
14	FC4.I have the resources necessary to use MyST (smartphone, computer, internet connection, microphone etc.)	Facilitating Conditions	Q6_3
15	SI1. My teachers or instructors think that I should use MyST	Social Influence	Q6_2
16	SI2. People who are important to me think that I should use MyST.	Social Influence	Q7_1
17	SI3. People from my university or school are encouraging the use of MyST.	Social Influence	Q8_4
18	Ht1. It has become my habit to learn languages with mobile apps	Habit	Q5_3
19	Ht2. I often learn language(s) in mobile or computer applications.	Habit	Q9_4
20	Ht3. I must use mobile or computer applications to learn languages.	Habit	Q8_2
21	HM1. Using MyST is fun.	Hedonic Motivation	Q5_4
22	HM2. Using MyST is enjoyable.	Hedonic Motivation	Q10_3
23	HM3. Using MyST is very entertaining.	Hedonic Motivation	Q9_5
24	At1. Using MyST is a good idea.	Attitude toward using technology	Q9_1
25	At2. MyST makes language learning more interesting.	Attitude toward using technology	Q7_2
26	At3. I like learning with MyST.	Attitude toward using technology	Q10_1