



Students' acceptance of MySpeechTrainer to improve spoken Academic English

Helmer Strik^{1,2,4,5}, Anna Ovchinnikova², Camilla Giannini³, Angela Pantazi⁴, Catia Cucchiari⁵

¹ Centre for Language Studies (CLS), Radboud University Nijmegen

² NovoLearning BV Nijmegen

³ Linguistics, Utrecht University

⁴ Linguistics, Radboud University Nijmegen

⁵ Centre for Language and Speech Technology (CLST), Radboud University Nijmegen

w.strik@let.ru.nl, anna@novo-learning.com, c.giannini@students.uu.nl,
angela.pantazi@student.ru.nl, c.cucchiari@let.ru.nl

Abstract

CALL systems with ASR-technology may provide a solution for the growing need to support students' English communication skills at universities, but so far developments in this direction have been limited. MySpeechTrainer is an ASR-based CALL system intended to help students improve their general Academic English oral communication skills, as well as subject-specific vocabulary. In this paper we report on a study that evaluated the acceptance of MySpeechTrainer from the perspective of its target users. To this end, the 'Unified Theory of Acceptance and Use of Technology' (UTAUT) model, which is frequently used in ICT research, was employed to analyze students' acceptance and use of the new system. Two groups of students from Nijmegen and Utrecht universities used and reviewed MySpeechTrainer (N=87). Results of statistical analysis indicated that Social Influence was one of the main predictors of students' readiness to use the system. Participants were generally positive about MySpeechTrainer and indicated the need for more supportive and informative feedback.

Index Terms: ASR-based CALL, Academic English, course specific content, UTAUT, Behavioral Intention, Social Influence

1. Introduction

The number of universities offering courses in English has increased significantly in recent years, as English has now become the lingua franca in higher education in many countries around the world. Consequently, there are more students who need to learn and use Academic English (AE), which is also often referred to as English for Academic Purposes (EAP). Even though most undergraduate students may have a proficient level in English, they still face many challenges in communication and vocabulary. First, it is important for them to acquire the specific language and terminology that is required in an academic environment and, second, it often turns out that they need additional practice to improve their English oral skills. However, practicing oral skills is time-consuming and

requires intensive practice and focused feedback, which cannot always be provided to a sufficient extent. For this reason, alternative solutions have been sought by developing Computer Assisted Language Learning (CALL) systems that employ Automatic Speech Recognition (ASR) technology. The implementation of ASR in CALL systems present many advantages for second language learning (Cucchiari & Strik, 2018). It provides additional practice, specific feedback on speaking skills, and the possibility for students to practice outside of the classroom.

Although several ASR-based products have been developed and commercialized (Foote, 2017; Golonka, Bowles, Frank, Richardson & Freynik, 2014; Savvani, 2018), there is still little information available about their performance and acceptability. Most of the evaluations conducted so far focus on the performance of the ASR technology and less on the real students' experience with the tool (Cucchiari & Strik, 2018; Hsu, 2016; Neri, Mich, Gerosa, & Giuliani, 2008). Furthermore, although previous studies already focused on the students' perception of various tools 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).

The present study aims at investigating an ASR-based CALL system from the perspective of its users. To the best of our knowledge so far no ASR-based CALL systems have been developed for the specific purpose of improving oral Academic English (AE). In the present study we address this knowledge gap by presenting a prototype application called MySpeechTrainer (MyST) and by investigating its potential from the perspective of its users.

MyST courses can be accessed via Novo App on a smartphone, tablet and PC (<https://www.novo-learning.com/>) (available on request). The application uses ASR technology to provide individual feedback on pronunciation, and it gives the possibility to create customized courses for language learning.

A range of exercises was created for the given research. Twenty-two activities for general AE were created based on the most frequently used academic words and phrases as suggested in Academic Word List (Coxhead, 2000). They were divided by topics of pronunciation, grammar, vocabulary and fixed expressions. Students of Radboud University Nijmegen only used these general AE exercises. At Utrecht University, participants were students that followed a Semantics and Pragmatics course or a Phonetics course, and subject specific activities were developed for these courses. MyST includes interactions such as pronunciation, multiple choice, drop down choice, drag and drop, dialogues etc.

To investigate the students' acceptance of the tool a revised version of the UTAUT2 model will be used (Venkatesh., Morris, Davis & Davis, 2003; Venkatesh, Thong & Xu, 2012).

2. Background

2.1. UTAUT

Many CALL studies lack a sound theoretical framework and suffer from an inconsistent research design (Chwo, Marek & Wu, 2018; Duman, Orhon & Gedik, 2015; Sharples & Kukulska-Hulme, 2010). A good option might be to use a validated theoretical framework in CALL research (Morton & Jack, 2010, Rapp & Kauf, 2018). One model that is widely used in ICT research, but not yet as widely in CALL, is UTAUT (Venkatesh et al., 2003). This is a model that allows predicting user acceptance and use of a system based on several behavioral constructs such as: Performance Expectancy (PE), Effort Expectancy (EE), Social Influence (SI) and Facilitating Conditions (FC), moderated by age, gender, experience and voluntariness of use.

PE was defined by Venkatesh et al. (2003) as "the degree to which an individual believes that using the system will help him or her to attain gains in job performance". EE is how easy users perceive the system to be. Social influence is "the degree to which an individual perceives that important others believe he or she should use the new system". Finally, FC is the extent to which the user is aware of infrastructure that supports the use of the system.

Venkatesh et al. (2012) proposed the UTAUT2 model with an improved variance that included three new constructs: Hedonic Motivation, Price Value and Habit to the UTAUT model). Hedonic motivation (HM) is defined as the "degree of fun or pleasure from using a system". Price value (PV) is defined as "users' perception of obtained value of the system in relation to its monetary cost". Habit (Ht) is the "behavior that is performed automatically or because of learning".

Attitude (At) has been a subject of contradictory opinions in technology acceptance research. Venkatesh et al. (2003) defined attitude toward using technology as "an individual's overall affective reaction to using a system". The UTAUT model does not include attitude because the researchers did not observe any significant relationship between attitude and intention in the statistical analyses. However, Dwivedi, Rana, Jeyaraj, Clement & Williams (2017) proposed an alternative model that brings attitude back into the acceptance model. In their meta-analysis of 162 studies, they determined that attitude

had direct effect on Behavioral Intention and actual use. They argued that attitude is an important predictor and moderator of both intention and use, and hence, should be included in research.

Although the UTAUT model was criticized for its complexity and a high number of disparate predictors (Bagozzi, 2007; van Raaij & Schepers, 2009), it has been used widely and effectively by numerous studies and in various fields (Williams, Rana & Dwivedi, 2015).

2.2. CALL and technology acceptance models

Relatively few studies used technology acceptance models in the context of CALL. Lai, Wang, Li & Hu (2016) investigated the influence of teacher support on students' acceptance of technology, comparing universities in Hong Kong and the U.S with 418 participants. SI and FC were found to be significant moderators of teacher's support. Tan (2013) studied the use of websites for learning English in Taiwan, collecting 176 responses. The study reported all four predictors to be significant, with PE and SI as the strongest predictors. Khechine, Lakhal, Pascot & Bytha (2014) collected 114 questionnaires to investigate acceptance and use of a webinar platform for various disciplines, including language learning, and found that PE, SI and FC were significant predictors.

Hsu (2012) investigated the use of the Moodle platform for computer-mediated communication amongst 47 students in Taiwan and found that PE, EE and SI were significant predictors. Liu (2013) also studied the acceptance of the Moodle platform by ESL students with 13 participants. and found PE and EE significant. Liu and Huang (2015) researched the use of Google Docs for translation practice with 27 students and found FC, SI and EE to be significant predictors. Doremalen, Boves, Colpaert, Cucchiari, and Strik (2016) developed and evaluated a prototype of an ASR-based CALL system named DISCO that provided pronunciation, syntax and morphology practice in Dutch L2, means of the responses were reported, and the conclusions were that the students were generally positive and liked the system.

Pynoo et al. (2011) investigated teachers' acceptance of a mandatory learning management system at a secondary school in Belgium (<https://www.smartschool.be>). As the tool was mandatory and already in use by 90 teachers, the researchers could compare the predicted intention with the actual use over the period of one year. Although re-using the same questionnaire multiple times might have caused respondent fatigue and a negative effect on the quality of responses (Lavrakas, 2008), the researchers could report about the changes in intention throughout the year. At the end of the year, PE appeared the main predictor for attitude, intention and actual use. EE predicted attitude, SI predicted intention, self-reported use and actual use. FC predicted actual use.

The present study uses the UTAUT model in an attempt to bridge the gap between technology and language learning, tackling the problem of inconsistent methodologies in CALL research and applying a well-tested model for evaluating a tool. Furthermore, this study aims to evaluate the acceptance of a tool with ASR technology.

3. Methods

3.1. The questionnaire

The questionnaire was created in Qualtrics (2005) and could be accessed through an internet link. It included 26 UTAUT statements, covering the constructs of PE, EE, SI, FC, At, Ht, HM, and demographic questions. 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, to 7 - Strongly agree). At the end of the questionnaire there was an open-ended text field where students could enter feedback, comments, etc.

3.2. Participants

Nijmegen

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. The students who agreed were instructed to complete 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. After data collection was completed, five randomly selected participants were rewarded with a coupon of 10 EUR value.

Utrecht

The participants in Utrecht were students who were following a Semantics and Pragmatics or a Phonetics course as part of their bachelor's degree.

MyST was presented by the teachers of the two courses 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 used MyST as an extra, non-graded assignment. At the end of the courses, the participants filled out the questionnaire. In total 36 students completed the questionnaire, 11 from the Semantics of Pragmatics course, and 25 from the Phonetics course.

3.3. Analyses

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. PLS-SEM was selected as it is adequate for small sample sizes, it focuses on prediction and interaction (Chin et al., 2003; Henseler et al., 2009). The software SmartPLS 3 was used for the application of PLS-SEM (<https://www.smartpls.com/>).

Table 1. Results for the constructs. For explanation see text.

Constructs	BI	PE	Ht	SI	At	EE	FC
Mean	3.71	4.93	3.29	3.87	5.12	5.81	6.21
Std.Dev.	1.56	1.18	1.51	1.18	1.09	0.89	0.78
VIF	-	3.14	1.46	1.42	3.72	2.65	1.59
Path Coef.	-	0.051	0.036	0.439	0.520	-0.201	0.077
p-value	-	0.356	0.353	0.000	0.001	0.039	0.210

4. Results

After excluding invalid questionnaires, 84 questionnaires remained (48 from Nijmegen, 36 from Utrecht; 24 male, 60 female). The age of the students ranged from 18-25, and 63% had an age between 19-21.

Because of multicollinearity, HM was removed, after which the variance inflation factor (VIF) values of all the remaining constructs remained below the threshold value of 5 indicating no collinearity issues (see Table 1). In addition, we tested the values of the outer loadings, internal consistency reliability focusing on composite reliability, Cronbach's Alpha, convergent validity using Average Variance Extracted (AVE), and discriminant validity using Heterotrait-Monotrait (HTMT) confidence interval scores. All values were adequate. The mean and standard deviation values of the constructs are presented in Table 1.

Next, we carried out an analysis of the structural model. The results in Table 1 show that At, EE and SI have a statistically significant impact on Behavioral Intention (BI) toward using MyST, with SI and At having the strongest magnitude on the relationship with BI. For the model, $R^2 = 0.606$, which can be considered moderate to substantial (Hair et al., 2014: 216), and thus the basic structure of UTAUT can explain 60.6 % of the variation in students' Behavioral Intention.

Forty-one participant provided comments and feedback in the open-ended text field at the end of the questionnaire. In general, they were positive about MyST (i.e. "I like that you have to talk out loud", "Helpful and clear concept. Good visuals and not complicated to use."), but some suggested that more feedback could be provided (i.e. "Feedback on the answers would be an addition.", "If I could not think of the correct answer, there was no way of finding out what the correct answer was, or why.").

5. Discussion

The research reported on in this paper addressed the development and evaluation of a CALL system, called MySpeechTrainer, that employs ASR technology to offer practice and feedback on Academic English oral skills. The novelty of the system relies on the combination of ASR technology within 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.

An important gap we identified in the literature is that few studies have investigated ASR-based CALL systems in realistic conditions with real users, and that many studies focus on ASR performance neglecting the users' perspective. Therefore, a second novelty of this study is that we chose to evaluate the system from the perspective of its users by adopting a technology acceptance model, i.e. UTAUT.

The mean values of the constructs related to the use of MyST, i.e. PE, EE, At, and FC, are high. Furthermore, the students made positive comments at the end of the questionnaire. These high mean values, the positive comments, and the significant relationship of At and EE with BI indicate not only that MyST is accepted by its target users, but also that they enjoy it and find it useful and easy to use. This is in contradiction with findings by Choi (2016), who suggested that even the most advanced neuro-linguistic programme technology is not capable of adequately examine learners' output and thus may cause frustration and reduce motivation. Since users of the ASR-based CALL system MyST were positive, MyST seems suitable to carry out more research.

An analysis of the answers to the open questions revealed that although the students were happy to receive feedback, they also pointed out that more informative feedback would be welcome. The feedback provided in MyST indicates to the learner that something was wrong with the utterance produced and it also pinpoints the error, including pronunciation errors at the phoneme level. Apparently, this was found to be insufficient by some users who preferred to receive additional information about what was precisely wrong and how they could improve on that point. This is an interesting issue from a pedagogical perspective and a challenging one from a speech technology perspective. Research on the role of corrective feedback in second language acquisition has shown that the effectiveness of feedback can vary depending on the type of error, but also on the learner. While for some learners feedback strategies like elicitations may be more effective because these force the learner to come up with the correct answer and are thus likely to leave more memory traces, there are also learners that prefer other forms of more elaborate feedback like metalinguistic feedback (Sheen & Ellis, 2011).

For speech technology providing detailed feedback on the nature of the error and on possible ways of eliminating its causes may be extremely challenging and can increase the chance of providing erroneous feedback, something that of course should always be avoided because extremely detrimental for the learner. However, addressing this challenge is interesting from a research point of view. In future studies it would be useful to include other forms of feedback that provide more information on how errors could be improved and to see how these are evaluated by different types of users.

Other than most of the studies that applied the UTAUT model in ICT research, our study supports theoretically and empirically the UTAUT reliability and validity in an academic educational context and more specifically in the CALL and ASR-technology research. Our research validates the extended UTAUT model presented by Dwivedi et al. (2017).

In the final model, significant results were obtained for SI, At, and EE. Most significant is SI, while path coefficients of SI

and At are higher than those for EE. We tried other types of statistical analysis, using other models, also for subsets of the data, and in all cases SI was the most significant construct. The findings for SI are in accordance with many previous studies on technology acceptance in language learning in spite of different characteristics of participants and different contexts of research (Hsu, 2016; Khechine, 2014; Tan, 2013). Significant results for At is in line with findings of Dwivedi et al. (2017), while it is in contradiction with previous research which did not include attitude, because researchers had not observed any significant relationship between attitude and intention in statistical analyses (Venkatesh et al., 2003; Venkatesh et al., 2012).

An interesting finding is that SI turned out to be the most significant factor for the two groups of users in Nijmegen and in Utrecht, although their practice differed from each other: short voluntary use vs. use of varying lengths. At the moment, CALL systems are often on a voluntary basis, not combined with standard courses at university, and are not recommended by teachers or other important persons. However, since SI appears to be the most important determinant, it seems that the acceptance and use of CALL systems could be increased by increasing SI.

Our research does not indicate the constructs of Performance Expectancy, Habit and Facilitating Conditions to be significant predictors of Behavioral Intention. Performance Expectancy may not be significant, because the users did not use MyST for a long time to be able to understand if this would help in the development of spoken Academic English. The study by Pynoo et al. (2011) indicates that there can be interesting changes in intention determinants of the use of a tool throughout a year. During the first month of use, Performance Expectancy and Social Influence were predictors of Behavioral Intention, while after some months of use Performance Expectancy remained as the sole significant predictor of Behavioral Intention.

6. Conclusions

Novel aspects of the current study are that we developed an ASR-based CALL system, MyST, to provide practice and feedback on oral Academic English, with course-specific content, and that we used the technology acceptance model UTAUT to evaluate the system. Users were positive about MyST. Tests carried out made clear that the final UTAUT model was valid and reliable; results showed significant results for SI, At, and EE. SI and At were more significant and path coefficients were higher than those for EE. A good strategy when using CALL systems might thus be to try to improve SI and EE, as improving At is probably more difficult. Furthermore, there is room for improvement, since users indicated that they would like to receive more detailed feedback.

7. Acknowledgements

We would like to thank all the students and teachers in Nijmegen and Utrecht that participated in our experiments.

8. References

- [1] Ayres, R. (2002). Learner attitudes towards the use of CALL. *Computer assisted language learning*, 15(3), 241-249.
- [2] Bagozzi, R.P. (2007). "The Legacy of the Technology Acceptance Model and a Proposal for a Paradigm Shift", *Journal of the Association for Information Systems*, 8 (4), pp. 244–254
- [3] Chin, W.W.(1998). The partial least squares approach to structural equation modeling. In: G. A.Marcoulides(Ed.), *Modern Methods for Business Research*. Mahwah, NJ: Lawrence Erlbaum Associates, 295–358.
- [4] Choi, I. C. (2016). Efficacy of an ICALL tutoring system and process-oriented corrective feedback. *Computer Assisted Language Learning*, 29(2), 334-364.
- [5] Chwo, G., Marek, M. & Wu V. (2018). Meta-analysis of MALL research and design. *System*. 74. 10.1016/j.system.2018.02.009.
- [6] Coxhead, Averil (2000) A New Academic Word List. *TESOL Quarterly*, 34(2): 213-238.
- [7] Cucchiarini, C., & Strik, H. (2018). Second Language Learners' Spoken Discourse: Practice and Corrective Feedback Through Automatic Speech Recognition. In *Smart Technologies: Breakthroughs in Research and Practice* (pp. 367-389). IGI Global.
- [8] Doremalen, J. van, Boves, L., Colpaert, J., Cucchiarini, C., & Strik, H. (2016). Evaluating automatic speech recognition-based language learning systems: a case study. *Computer Assisted Language Learning*, 29(4), 833–851. <https://doi.org/10.1080/09588221.2016.1167090>
- [9] Duman, G., Orhon, G., & Gedik, N. (2015). Research trends in mobile assisted language learning from 2000 to 2012. *ReCALL*, 27(2), 197–216. <https://doi.org/10.1017/S0958344014000287>
- [10] 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*, 1-16.0
- [11] Foote, J. A., & McDonough, K. (2017). Using shadowing with mobile technology to improve L2 pronunciation. *Journal of Second Language Pronunciation*, 3(1), 34-56.
- [12] 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.
- [13] 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.
- [14] 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.
- [15] Hsu, H. (2012). The Acceptance of Moodle: An Empirical Study Based on UTAUT. *Creative Education*, 03(08), 44–46. <https://doi.org/10.4236/ce.2012.38B010>
- [16] 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.
- [17] Khechine, H., Lakhali, 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.
- [18] 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, 676–688. <https://doi.org/10.1016/j.chb.2016.04.039>
- [19] Liu, J. (2013). E-learning in English classroom: Investigating factors impacting on ESL (English as Second Language) college students' acceptance and use of the Modular Object-Oriented Dynamic Learning Environment (Moodle). 114.
- [20] 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.
- [21] McCrocklin, S. (2019). ASR-based dictation practice for second language pronunciation improvement. *Journal of Second Language Pronunciation*, 5(1), 98-118.
- [22] Morton, H., & Jack, M. (2010). Speech interactive computer-assisted language learning: a cross-cultural evaluation. *Computer Assisted Language Learning*, 23(4), 295–319. <https://doi.org/10.1080/09588221.2010.493524>
- [23] 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.
- [24] Pynoo, B., Devolder, P., Tondeur, J., Van Braak, J., Duyck, W., & Duyck, P. (2011). Predicting secondary school teachers' acceptance and use of a digital learning environment: A cross-sectional study. *Computers in Human behavior*, 27(1), 568-575.
- [25] Qualtrics (2005). Qualtrics, Provo, Utah, USA. Retrieved from <http://www.qualtrics.com>.
- [26] Rapp, C., & Kauf, P. (2018). Scaling Academic Writing Instruction: Evaluation of a Scaffolding Tool (Thesis Writer). *International Journal of Artificial Intelligence in Education*, 1–26. <https://doi.org/10.1007/s40593-017-0162-z>
- [27] Savvani, S. 2018. State-of-the-Art Duolingo Features and Applications. In *International Conference on Interactive Collaborative Learning* (pp. 139-148). Springer, Cham.
- [28] Sharples, M., & Kukuluska-Hulme, A. (2010). Learning using mobile and handheld devices. Retrieved from http://repository.alt.ac.uk/2301/1/Learning_using_mobile_and_handheld_devices.docx.
- [29] Sheen, Y., & Ellis, R. (2011). Corrective feedback in language teaching. *Handbook of research in second language teaching and learning*, 2, 593-610.
- [30] 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.
- [31] Raaij, E. M. van, & Schepers, J. J. L. (2008). The acceptance and use of a virtual learning environment in China. *Computers & Education*, 50(3), 838–852. <https://doi.org/10.1016/j.compedu.2006.09.001>
- [32] Venkatesh, V. and H. Bala, (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences*, 39(2), 273-315.
- [33] Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478.
- [34] 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.
- [35] 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.
- [36] 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>