IberSPEECH 2021
24-25 March 2021, Valladolid, Spain

Design and Evaluation of Mobile Computer-Assisted Pronunciation Training Tools for Second Language Learning: a Ph.D. Thesis Overview

Cristian Tejedor-García[1], *Valentín Cardeñoso-Payo*[1], *David Escudero-Mancebo*[1]

[1]ECA-SIMM Research Group, Department of Computer Science, University of Valladolid, Spain

cristian@infor.uva.es

Abstract

Recent advances on speech technologies (automatic speech recognition, ASR, and text-to-speech, TTS, synthesis) have led to their integration in computer-assisted pronunciation training (CAPT) tools. However, pronunciation is an area of teaching that has not been developed enough since there is scarce empirical evidence assessing the effectiveness of CAPT tools and games that include ASR/TTS. In this manuscript, we summarize the findings presented in Cristian Tejedor-García's Ph.D. Thesis (University of Valladolid, 2020). In particular, this dissertation addresses the design and validation of an innovative CAPT system for smart devices for training second language (L2) pronunciation at the segmental level with a specific set of methodological choices, such as the inclusion of ASR/TTS technologies with minimal pairs, learner's native–foreign language connection, a training cycle of exposure–perception–production, and individual/social approaches. The experimental research conducted applying these methodological choices with real users validates the efficiency of the CAPT prototypes developed for the four main experiments of this dissertation about English and Spanish as L2. We were able to accurately measure the relative pronunciation improvement of the individuals who trained with them. Expert raters on phonetics' subjective scores and CAPT's objective scores showed a strong correlation, being useful in the future to be able to assess a large amount of data and reducing human costs.

Index Terms: Computer-assisted pronunciation training (CAPT), second language (L2) pronunciation, automatic speech recognition (ASR), text-to-speech (TTS), autonomous learning, automatic assessment tools, learning environments, mobile learning game, minimal pairs

1. Introduction

Computer-assisted pronunciation training (CAPT) tools are ideal and attractive for language learning since they provide learner autonomy and individualized and tireless instruction [1]. When these tools are well designed, it is expected to provide learners with more pronunciation practice than in traditional courses [2], real time and consistent feedback [3], and the automatic measurement of pronunciation quality [4].

In this manuscript, we present a summary of the main findings of the Ph.D. Thesis defended by Cristian Tejedor-García on the 30th of September, 2020 at University of Valladolid (Spain), *cum laude* and internal mentions. We designed and evaluated different CAPT experiments in which the main second languages (L2) were English and Spanish. We also shed light about how far we are with respect to the ideal CAPT scenario presented in the previous paragraph. Interested readers can find the numerical and statistical results of the experimentation in the complete dissertation manuscript [5].

This work was supported in part by the Ministerio de Economía y Empresa (MINECO), in part by the European Regional Development Fund FEDER under Grant TIN2014-59852-R and Grant TIN2017-88858-C2-1-R, in part by the Consejería de Educación de la Junta de Castilla y León under Grant VA050G18 and Grant VA145U14, and in part by the University of Valladolid (Ph.D. Research Grant 2015 and MOVILIDAD DOCTORANDOS UVa 2019).

1.1. Motivation

There are approximately 2.7 thousand million L2 speakers worldwide [6] and e-learning is an attractive and emerging option as a complement for one-to-one tutoring and classroom courses [7, 8]. Mobile learning applications and games are appearing for language learning for two reasons. First, they offer anytime, anywhere, and high intensity training [9]; and the quality of automatic speech recognition (ASR) [10] and text-to-speech (TTS) synthesis modules [11] that they include has considerably been improved in the last decade. However, there have been few attempts to empirically measure the effectiveness of mobile CAPT tools with ASR and TTS systems [2].

The reasons why it is not easy to measure the effectiveness of CAPT tools are mainly five. First, the similarities and differences of the native language (L1) and L2 must be taken into account; the training activities should be adapted to each learner; and the feedback provided could be insufficient or too difficult to understand by the learners [12]. Also, the automatic measurement of the quality of learner's pronunciation is not trivial and the ASR/TTS modules selected must be tested previously. Fourth, game elements and interaction with other learners might deviate from the goal of pronunciation improvement or discourage learners [13]. Finally, although multidisciplinarity is in most cases an advantage, CAPT needs to deal with some issues of communication between experts from different areas since they focus on different aspects when evaluating [1]. Thus, the current challenge is to carefully design and adapt an effective CAPT system with not only such speech technologies, but also with a training methodology, a pronunciation improvement assessment, and a corrective feedback strategy, according to learner's L1 and L2.

1.2. Research Objectives and Questions

The main objective of this thesis is to design and evaluate a CAPT tool for smart devices which incorporates current ASR and TTS technologies; helping learners to work autonomously, at their own pace, and with the possibility of providing real-time feedback. It is divided into four specific research objectives:

RO1. To analyze and define a set of activities, protocols, and motivational elements for the improvement of L2 pronunciation with a CAPT system which integrates ASR and TTS technologies.

RO2. To select the most appropriate metrics for the assessment of the learner's pronunciation level.

10.21437/IberSPEECH.2021-35

RO3. To design a semi-automatic method supervised by experts for obtaining a specific set of minimal pairs adapted to L2 pronunciation problems, according to the learner's L1 and to the limitations of the ASR and TTS technologies.

RO4. To select and design a CAPT system with current ASR and TTS technologies that provides an individualized feedback to the learner for improving L2 pronunciation.

In order to carry out the experimental procedure, three research questions were identified to validate the research objectives, categorized by topics. The first topic was related to the feasibility of current speech technology (ASR and TTS systems) integration in CAPT tools:

RQ1. Can current ASR and TTS systems be successfully used in a non-obstructive way in the CAPT tool developed?

Issue 1.1. Can current ASR and TTS systems help to assess different groups of learners according to their L2 pronunciation level in the CAPT tool developed?

The second topic referred to the implications of the training methodology with CAPT tools in learner's pronunciation improvement:

RQ2. To what extent can methodologically sensitive design issues, such as the use of exercises based on minimal pairs within the training activities cycle proposed in the CAPT tool developed affect user's pronunciation improvement?

Issue 2.1. Can a relative improvement in the learner's pronunciation be assessed after using the CAPT tool?

Issue 2.2. If any, is there a relevant pronunciation improvement from a quantitative point of view?

Issue 2.3. Does the tool reveal what the real difficulties of the users are (most difficult sounds/training activities)?

Finally, the last research question aimed at answering how game elements and social approaches affect learner's implication in pronunciation training with CAPT tools:

RQ3. To what extent can gamified versions of the tool affect user's motivation, performance, and learning?

1.3. Research Methodology

An experimental research [14] was conducted for the whole experimentation process to accomplish the objectives and give answers to the research questions proposed in this thesis, with a multidisciplinary group of researchers and experts.

Five phases were followed in each experimental iteration. First, the research problem was identified. The process started by clearly identifying the problems that will be addressed during the research process, examining the existing solutions in the state-of-the-art, and considering what possible methods will affect such solutions. Second, it was carefully planned and devised the experimental research study to test the research objectives and questions: (1) the participants were selected, that is, target population, enrollment rules, sample size, and groups; (2) different metrics were defined to measure the research variables from the data results gathered from the instruments; (3) the assessment protocol was defined, the research variables were measured before, during, and after performing the training activities; and (4) a CAPT tool was developed for each experiment of the thesis. Third, the experiment was conducted. At the beginning, the participants' groups were established. Then, each user performed the activities defined for her/his group in the previous phase, and the experimental data related to the variables of the study was collected with specific instruments for each experiment. Fourth, the data gathered was analyzed indicating the relevant indicators in order to corroborate whether the experiment was successful. Finally, the most relevant results were shared and published in scientific journals and conferences by means of articles, abstracts, show and tell demonstrations, and presentations.

2. Experimental Procedure

Given the complexity of the task, it was prudent to address it in an incremental and evolutive approach to facilitate that the results of each one of the experiments could be analyzed to change and refine the design of the next one (see Figure 1).

The evolution of the prototypes was aligned with two main focuses for the different methodologies followed to address the L2 pronunciation training through four experiments and six prototypes[1]. The first focus, *Game* (at the bottom of the figure), proposed incorporating gamification elements and social strategies; whereas the second focus, *Guided*, followed an individual approach with guided training instructions.

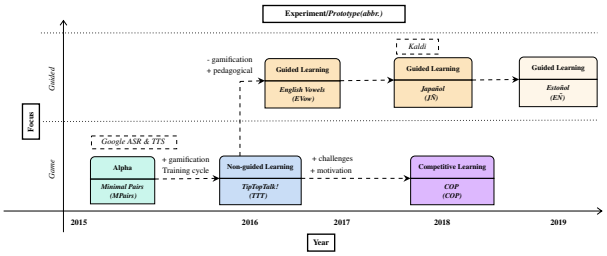

Figure 1: *Evolution diagram of the experiments and prototypes.*

2.1. Game-based Experiments

The first experiment was called **Alpha**, and the prototype developed was **Minimal Pairs**. It was a proof of concept, a starting point for checking the viability of including state-of-the-art general–purpose ASR and TTS technologies (Google) in an L2 pronunciation training protocol with minimal pairs (two similar sounding words that differ in only one phonological element and have distinct meanings); and their possible weaknesses and limitations. We also wanted to find reliable metrics to assess pronunciation tasks (see an example in Figure 2); and gather opinions from the users of a focus group session for future experiments

The number of uses of the speech technologies, success rates, and the position of the target word in the ASR hypotheses list (RO1), proved to be valid metrics (RO2) to relate the performance results with minimal pairs activities (RO3) to the proficiency level declared of each one of the participants of the three different groups (RQ1): natives, non-natives advanced-level learners, and non-natives beginner-level ones. The use of TTS by non-natives was isolated and voluntary; though no significant improvements in success rates were obtained after its use. Besides, some limitations in ASR technology were found, as native learners did not successfully complete all production activities. Finally, the opinions in the focus group session about the tool suggested new activities, feedback techniques, and motivational elements to consider for future experiments.

After checking the viability of the application of commercial ASR/TTS speech technologies and minimal pairs for a L2 pronunciation training protocol in the first experiment, we

[1]The ongoing Estoñol prototype started at the end of this Ph.D. Thesis and is still active with the collaboration of the University of Tartu, Estonia.

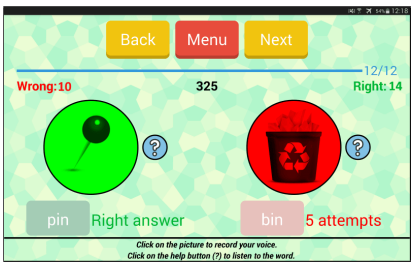

Figure 2: *Minimal Pairs prototype CAPT tool's interface [15].*

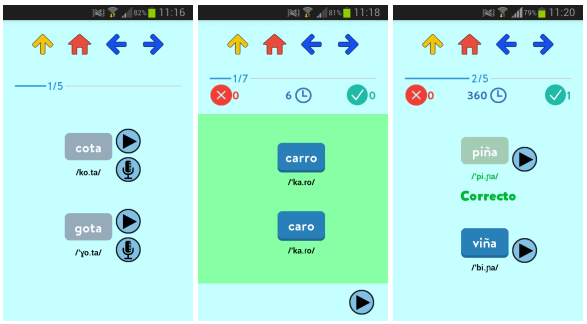

Figure 3: *TipTopTalk! prototype CAPT tool's interfaces [16].*

wanted to assess the possible pronunciation improvement after training for a longer period of time. Following the recommendations of the users, we included more training activities in the second experiment, **Non-guided Learning**: exposure, discrimination, and production tasks. We also included more gamification elements to motivate players to keep on playing. so we designed an individual competition **TipTopTalk!** in which learners played with the tool (see Figure 3), installed in their own smart devices.

The activities of perception (with TTS) were the most trained ones (RO1, RQ1), which led to a general improvement in the ability of the players (RO2, RO3, RO4, RQ2). They tended to train the easiest activities for achieving positive results. However, despite the introduction of gamification elements (RQ3) and the improvement in discrimination skills, a production improvement stagnation and a loss of interest was detected in the most proficient players. Finally, the answers to questionnaires revealed good usability results and the requirement of new feedback techniques in future work.

The training approach in the previous experiments was individual in all cases. Participants performed different activities with the tool without sharing tasks. At this point, we considered the possibility of including not only game elements, but also challenges between participants to improve their motivation and performance, improving the results of participation registered in the prototype, TipTopTalk! In order to do so, we designed a second version of the game, called **COP**, with common game rules in which each player participated in challenges with others via a mobile application (see Figure 4). In these challenges with ASR, TTS and minimal pairs, the players achieved points, and climbed up a leaderboard, while we gathered a large amount of speech and behavior data. We wanted to measure the effects of an intensive use of the game (high participation in challenges) on the most active user's motivation, performance, and learning improvement.

The lessons learned of this last game-based prototype, COP,

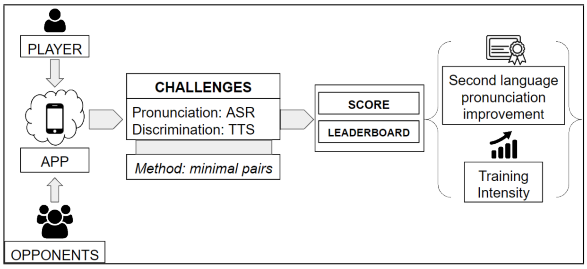

Figure 4: *COP prototype's conceptual view [17].*

were three. First, results showed an intensive practice (RQ1, RO1, RO2, RO3, RO4) supported by a significant quantity of data gathered, activities, and game days (RQ3). The most active and motivated players achieved significant pronunciation improvement results (RQ2). These outcomes were supported by the answers to the questionnaires and focus groups.

2.2. Guided-based Experiments

The second focus of the experimentation started with the **Guided Learning** experiment and **English Vowels** prototype. Both the stagnation detected in the second prototype TipTopTalk! on the most proficient learners in production improvement and their interest on performing the easiest activities, motivated this new focus of research, more pedagogical, guided, and individualized. The main objective was to train users' pronunciation by guiding them through a CAPT tool with personalized and more accurate feedback, in a controlled training protocol based on user's results and with a pre/post test strategy with minimal pairs. We included an in-classroom group to measure the proficiency improvement without a tool but with the same training activities with a teacher. Finally, we also wanted to measure the correlation between human ratings and ASR ones.

Then, we wanted to take a step further the English Vowels prototype, so we looked for alternatives to the general–purpose Google ASR system to assess a greater quantity of utterances, developing an in–house ASR system with Kaldi to address the question of how these general and specific–purpose ASR systems can deal with the assessment of minimal pairs in the field of CAPT. As a consequence, we designed **Japañol**, a CAPT tool for training Spanish pronunciation for Japanese native learners. Figure 5 shows the main training steps in the Japañol tool.

Results of both prototypes showed an intensive and effective practice (RQ1) that led to a significant improvement in pronunciation skills (RQ2). This improvement was higher in the CAPT tool group (RO1, RO3). In addition, the feedback offered in the training turned out to be effective (RO4), since those who followed it obtained better results than those who did not. Also we found strong correlations between the scores of the CAPT tool, human raters' scores, and ASR ones in the post-test (RO2). Finally, the tailored Kaldi–based ASR developed for the Japañol prototype showed that it can be as useful as a general–purpose ASR for assessing minimal pairs in CAPT tools (RO4) opening the possibility of designing a detailed study of pronunciation errors and results for future projects.

3. Conclusions

The main contributions of this Ph.D. Thesis are the following. First, we concluded that ASR and TTS technologies were ef-

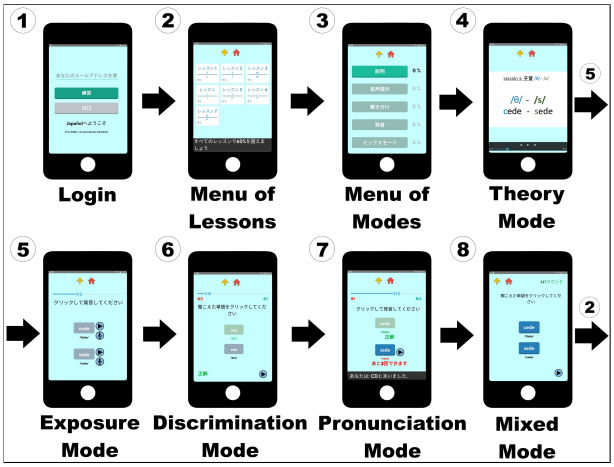

Figure 5: *Standard flow to complete a lesson in Japañol [18].*

fectively incorporated in a non-obstructive way to the L2 CAPT tools developed. Also, they proved to be a very useful didactic instrument that can be used complementary with other forms of second language acquisition.

Regarding training methodology, the versions of the tool designed and validated, and the methodological choices that they implemented, allowed us to measure the pronunciation improvement of the individuals who trained with them. Also, different corrective feedback techniques proved to be useful to overcome the proposed training activities. Our automatic data gathering method allowed us to provide specific feedback to users instantly and to analyze all results at the end of the experiments. Moreover, the minimal pair lists were elaborated with a novel semi-automatic protocol, taking into account learner's L1 and L2 and the specific ASR and TTS technologies [19].

We reported on positive results from both subjective and objective techniques for assessing pronunciation improvement during and after using the tools. The CAPT values highly correlated ($r > 0.8$) to expert evaluation. Results reported from the guided-based prototypes English Vowels and Japañol were very promising since the students achieved significant pronunciation improvement values (around one point on a ten-point grading scale), being higher in those who worked with the tool. In the COP prototype, the most active learners achieved the best pronunciation improvement and motivation results.

Finally, the game elements included in the tools proved capable of motivating learners to keep on playing on their own. The COP challenges proved to be a positive motivational factor since the most active learners made an intensive use of the game and significantly improved their L2 pronunciation along time.

As main limitations of this work, we assume that a closer comparative research between the activity and results of the tools and those of human-led instruction, such as video-taping or log monitoring might have helped us to obtain a more detailed knowledge on the possibilities of CAPT. However, our goal was not comparing autonomous systems with human instruction, but assessing the range of improvement procured by the use of the tools. Finally, although the focus of this thesis has been pronunciation improvement at the segmental level, its combination with the suprasegmental level and other focuses, such as intelligibility or fluency, must be also considered for acquiring a complete pronunciation competence.

4. Scientific Contributions

The scientific impact of this Ph.D. Thesis was supported by sixteen publications at the time of the dissertation presentation. In particular, game-based results of the Alpha experiment have been partially published in [15]. Performance–related results of the TipTopTalk! prototype have been partially published in [16, 20, 21]; whereas results related to the gamification elements included in the CAPT tool have been published in [22, 23]. Regarding the latest game-based prototype, COP, the main results about performance, motivation, and pronunciation improvement have been published in [17] (Journal Citation Reports, JCR Q1).

On the other hand, guided-based experimental results of the English Vowels prototype have been partially published in [24] (JCR Q2) and in [25, 26]. Results of pronunciation improvement and tool description of the Japañol prototype have been partially published in [18, 27, 28]. Other contributions related to this prototype have been partially published in [29, 19]. Also, first results of the latest and ongoing project, Estoñol, have been published in [30] (JCR Q2).

Finally, as a consequence of this work, we have attended nine international conferences and workshops, registered one software program (intellectual property), carried out two predoctoral research stays, obtained two predoctoral fellowships, participated in three research projects, achieved three awards, and developed six CAPT software applications.

5. Acknowledgments

Special thanks to Dr. Valentín Cardeñoso-Payo and Dr. David Escudero-Mancebo, the supervisors of this Ph.D. Thesis; and Dr. César González-Ferreras, Enrique Cámara-Arenas, and Dr. Mario Corrales-Astorgano from the ECA-SIMM research group. We would also like to thank Dr. María Jesús Rodríguez-Triana and the members of the Centre of Excellence in Educational Innovation at the University of Tallinn, Estonia; Mrs. Katrin Leppik and the members of the Institute of Estonian and General Linguistics (University of Tartu, Estonia); Dr. María J. Machuca and the members of Servei de Tractament de la Parla i del So (Autonomous University of Barcelona); and Mr. Takuya Kimura from the Spanish Language and Literature Department (University of Seisen). We truly appreciate the support of the colleagues and researchers of the Department of Computer Science at University of Valladolid, and the financial support of this University for carrying out this work. Also, our sincere thanks and appreciation for the reviewers and the examination board of this Ph.D. Thesis. Finally, we would like to thank the participation of more than five hundred learners in the experiments with the help of companies, language learning centers, and other researchers from all around the world.

6. References

[1] O'Brien *et al.*, "Directions for the future of technology in pronunciation research and teaching," *J. Second Lang. Pronunciation*, vol. 4, no. 2, pp. 182–207, Feb. 2018. [Online]. Available: https://doi.org/10.1075/jslp.17001.obr

[2] R. I. Thomson and T. M. Derwing, "The effectiveness of L2 pronunciation instruction: A narrative review," *Appl. Linguistics*, vol. 36, no. 3, pp. 326–344, Jul. 2015. [Online]. Available: https://doi.org/10.1093/applin/amu076

[3] B. Penning de Vries, C. Cucchiarini, H. Strik, and R. van Hout, "The Role of Corrective Feedback in Second Language Learning: New Research Possibilities by Combining CALL and Speech

Technology," in *Proc. SLATE*, Tokyo, Japan, Sep. 22–24, 2010, pp. 125–130.

[4] G. Seed and J. Xu, "Integrating technology with language assessment: Automated speaking assessment," in *Proc. ALTE*, Bologna, Italy, May 3–5, 2017, pp. 286–291.

[5] C. Tejedor-García, "Design and Evaluation of Mobile Computer-Assisted Pronunciation Training Tools for Second Language Learning," Ph.D. dissertation, Dept. Comput. Sci., Univ. Valladolid, Valladolid, Spain, Sep. 2020. [Online]. Available: https://doi.org/10.35376/10324/43663

[6] Wikipedia contributors, "List of languages by total number of speakers — Wikipedia, the free encyclopedia," https://en.wikipedia.org/w/index.php?title=List_of_languages_by_total_number_of_speakers, 2020, [Online; accessed 30-December-2020].

[7] K. Beatty, *Teaching & Researching: Computer-Assisted Language Learning*. Routledge, 2013. [Online]. Available: https://doi.org/10.4324/9781315833774

[8] H. D. Brown and H. Lee, *Teaching by principles: An interactive approach to language pedagogy*. New York, NY, USA: Pearson Education, 2015, vol. 1. [Online]. Available: https://doi.org/10.2307/3587655

[9] Cheon, Jongpil and Lee, Sangno and Crooks, Steven M. and Song, Jaeki, "An investigation of mobile learning readiness in higher education based on the theory of planned behavior," *Comput. Educ.*, vol. 59, no. 3, pp. 1054–1064, Nov. 2012. [Online]. Available: https://doi.org/10.1016/j.compedu.2012.04.015

[10] M. Meeker, "Internet trends 2017," may 2017, Kleiner Perkins, Los Angeles, CA, USA, Rep. [Online]. Available: https://www.bondcap.com/report/it17.

[11] J. Shen, R. Pang, R. J. Weiss, M. Schuster, N. Jaitly, Z. Yang, Z. Chen, Y. Zhang, Y. Wang, R. Skerry-Ryan *et al.*, "Natural TTS synthesis by conditioning WaveNet on mel spectrogram predictions," in *Proc. ICASSP*, Alberta, AB, Canada, Apr. 15–20, 2018, pp. 4779–4783. [Online]. Available: https://doi.org/10.1109/ICASSP.2018.8461368

[12] A. Neri, C. Cucchiarini, and H. Strik, "The effectiveness of computer-based speech corrective feedback for improving segmental quality in L2–Dutch," *ReCALL*, vol. 20, no. 2, pp. 225–243, May 2008. [Online]. Available: https://doi.org/10.1017/S0958344008000724

[13] T. Greitemeyer and D. O. Mügge, "Video Games Do Affect Social Outcomes: A Meta-Analytic Review of the Effects of Violent and Prosocial Video Game Play," *Personality Social Psychol. Bulletin*, vol. 40, no. 5, pp. 578–589, 2014. [Online]. Available: https://doi.org/10.1177/0146167213520459

[14] D. T. Campbell and J. C. Stanley, *Experimental and quasi-experimental designs for research*. Ravenio Books, 2015.

[15] D. Escudero-Mancebo, E. Cámara-Arenas, C. Tejedor-García, C. González-Ferreras, and V. Cardeñoso-Payo, "Implementation and test of a serious game based on minimal pairs for pronunciation training," in *Proc. SLaTE*, Leipzig, Germany, Sep. 4–5, 2015, pp. 125–130.

[16] C. Tejedor-García, V. Cardeñoso-Payo, E. Cámara-Arenas, C. González-Ferreras, and D. Escudero-Mancebo, "Measuring pronunciation improvement in users of CAPT tool TipTopTalk!" in *Proc. Interspeech*, San Francisco, CA, USA, Sep. 8–12, 2016, pp. 1178–1179.

[17] C. Tejedor-García, D. Escudero-Mancebo, V. Cardeñoso-Payo, and C. González-Ferreras, "Using challenges to enhance a learning game for pronunciation training of English as a second language," *IEEE Access*, vol. 8, no. 1, pp. 74 250–74 266, Apr. 2020. [Online]. Available: https://doi.org/10.1109/ACCESS.2020.2988406

[18] C. Tejedor-García, V. Cardeñoso-Payo, M. J. Machuca, D. Escudero-Mancebo, A. Ríos, and T. Kimura, "Improving Pronunciation of Spanish as a Foreign Language for L1 Japanese Speakers with Japañol CAPT Tool," in *Proc. IberSPEECH*, Barcelona, Spain, Nov. 21–23, 2018, pp. 97–101. [Online]. Available: https://doi.org/10.21437/IberSPEECH.2018-21

[19] C. Tejedor-García and D. Escudero-Mancebo, "Uso de pares mínimos en herramientas para la práctica de la pronunciación del español como lengua extranjera," *Revista de la Asociación Europea de Profesores de Español. El español por el mundo*, vol. 1, no. 1, pp. 355–363, Jan. 2018.

[20] C. Tejedor-García, D. Escudero-Mancebo, E. Cámara-Arenas, C. González-Ferreras, and V. Cardeñoso-Payo, "TipTopTalk! mobile application for speech training using minimal pairs and gamification," in *Proc. IberSPEECH*, Lisbon, Portugal, Nov. 23–25, 2016, pp. 425–432.

[21] C. Tejedor-García, D. Escudero-Mancebo, C. González-Ferreras, E. Cámara-Arenas, and V. Cardeñoso-Payo, "Improving L2 production with a gamified computer-assisted pronunciation training tool, TipTopTalk!" in *Proc. IberSPEECH*, Lisbon, Portugal, Nov. 23–25, 2016, pp. 177–186.

[22] C. Tejedor-García, V. Cardeñoso-Payo, E. Cámara-Arenas, C. González-Ferreras, and D. Escudero-Mancebo, "Playing around minimal pairs to improve pronunciation training," in *Proc. IFCASL*, ser. Feedback in Pronunciation Training Workshop, Saarland, Germany, Nov. 5–6, 2015.

[23] A. Rauber, C. Tejedor-García, V. Cardeñoso-Payo, E. Cámara-Arenas, C. González-Ferreras, D. Escudero-Mancebo, and A. Rato, "TipTopTalk!: A game to improve the perception and production of L2 sounds," in *Abstr. New Sounds 8th Int. Conf. Second Lang. Speech*, Aarhus Univ., Aarhus, Denmark, Jun. 10–12, 2016, p. 160.

[24] C. Tejedor-García, D. Escudero-Mancebo, E. Cámara-Arenas, C. González-Ferreras, and V. Cardeñoso-Payo, "Assessing pronunciation improvement in students of English using a controlled computer-assisted pronunciation tool," *IEEE Trans. Learn. Technol.*, vol. 13, no. 2, pp. 269–282, Mar. 2020. [Online]. Available: https://doi.org/10.1109/TLT.2020.2980261

[25] C. Tejedor-Garcia, V. Cardeñoso-Payo, and D. Escudero-Mancebo, "Design and evaluation of two mobile computer-assisted pronunciation training tools to favor autonomous pronunciation training of english as a foreign language," in *Proc. EDULEARN20*. IATED, 6-7 July, 2020 2020, pp. 7639–7646. [Online]. Available: http://dx.doi.org/10.21125/edulearn.2020.1936

[26] C. Tejedor-García, D. Escudero-Mancebo, C. González-Ferreras, E. Cámara-Arenas, and V. Cardeñoso-Payo, "Evaluating the efficiency of synthetic voice for providing corrective feedback in a pronunciation training tool based on minimal pairs," in *Proc. SLaTE*, Stockholm, Sweden, Aug. 25–26, 2017, pp. 26–30. [Online]. Available: https://doi.org/10.21437/SLaTE.2017-5

[27] C. Tejedor-García, V. Cardeñoso-Payo, and D. Escudero-Mancebo, "Japañol: a mobile application to help improving Spanish pronunciation by Japanese native speakers," in *Proc. IberSPEECH*, Barcelona, Spain, Nov. 2018, pp. 157–158.

[28] T. Kimura, C. Tejedor-García, V. Cardeñoso-Payo, M. J. Machuca, D. Escudero-Mancebo, and A. Ríos, "Japañol, a Computer Assisted Pronunciation Tool for Japanese Students of Spanish Based on Minimal Pairs," in *Abstr. 2nd Int. Symp. Appl. Phonetics*, Aizu, Japan, Sep. 21–23, 2018.

[29] C. Tejedor-García, "Design and Evaluation of a Mobile Application for Second Language Pronunciation Training based on Minimal Pairs," in *Proc. SEPLN 2018*, Seville, Spain, Sep. 21–23, 2018, pp. 7–11. [Online]. Available: http://ceur-ws.org/Vol-2251/paper2.pdf

[30] K. Leppik and C. Tejedor-García, "Estoñol, a computer-assisted pronunciation training tool for Spanish L1 speakers to improve the pronunciation and perception of Estonian vowels," *J. Estonian Finno-Ugric Linguistics (ESUKA – JEFUL)*, vol. 10, no. 1, pp. 89–104, Nov. 2019. [Online]. Available: https://doi.org/10.12697/jeful.2019.10.1.05