# Japañol: a mobile application to help improving Spanish pronunciation by Japanese native speakers

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

Grupo de Investigación ECA-SIMM. Universidad de Valladolid

valentin.cardenoso@uva.es

## Abstract

In this document, we describe the mobile application Japañol[1], a learning tool which helps pronunciation training of Spanish as a foreign language (L2) at a segmental level. The tool has been specifically designed to be used by native Japanese people, and implies a branch of a previous CAPT gamified tool TipTopTalk!. In this case, a predefined cycle of actions related to exposure, discrimination and production is presented to the user, always under the minimal-pairs approach to pronunciation training. It incorporates freely available ASR and TTS and provides feedback to the user by means of short video tutorials, to reinforce learning progression.

**Index Terms**: computer-assisted pronunciation training, speech recognition, human-computer interaction, computational para-linguistics.

## 1. Introduction

The way we teach and learn foreign languages is adapting to the technologies. The development of new modes to engage people in learning, as Computer-Assisted Pronunciation Training (CAPT), Computer-Assisted Language Learning (CALL) and Mobile-Assisted Language Learning (MALL), allow improving linguistic skills anytime and anywhere [1]. Also, these systems can give useful information on how learners perform and improve their pronunciation [2]. While there are many software tools that rely on speech technologies for providing to users L2 pronunciation training in the field of Computer Assisted Pronunciation Training (CAPT), Japañol [3], distinctively incorporates a well designed cycle of all the relevant activities related to pronunciation training: exposure, discrimination, production and mixed mode.

This application represents an evolution of previous serious games [4, 5, 6] designed for pronunciation training of L2 by non-native. All of them rely on the minimal pairs methodology [7] and are within the context of research projects related to the development and testing of software tools and games for foreign language learning (TIN2014-59852-R and VA050G18).

Previous versions were based on the free selection by users of exposure, discrimination and production tasks of, mainly, English or Spanish minimal-pairs, in order to get achievements and increase points in leader-boards. With that approach, we were able to assess user's pronunciation level in a L2 [8]. We have also analyzed how the introduction of corrective feedback [9, 10] increased pronunciation improvement among users after the first stages of use.
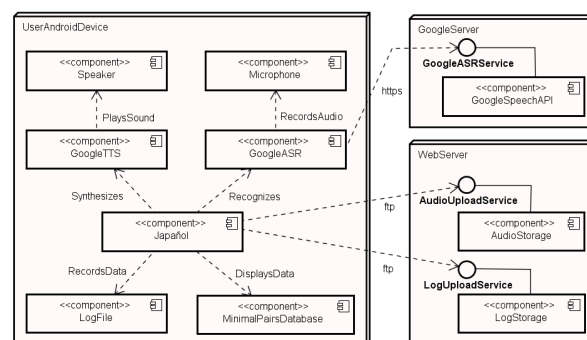
Figure 1: *Components of the CAPT system.*

While the freedom of movement on game-oriented tools leads users to maximize their score by repeating those tasks they found easy, the continued use of the tool seemed to generate stagnation. Complementary, learning-oriented tools should focus on users' difficulties, offering guided and corrective feedback and achieving better effectiveness and efficiency pedagogical results [11]. This motivated this alternate version of the tool, which provides a fixed cycle of well known and balanced learning activities, and which has been applied initially to an experiment which results are presented in this same conference [12].

## 2. System's description

Figure 1 shows the architecture diagram of components in Japañol. It is a native Android application built from scratch and can be run using low-cost resources available at language laboratories such as computers, tablets, speakers and microphones. It uses Google speech technology, such as Text-To-Speech offline tool and Automatic Speech Recognition web service for Android, that offers a N-best list of probable results for each utterance. Japañol keeps record of chronological events and results of the users with the system in log files. Both audio and log files are stored in a server, through a set of web services, in order to be later analyzed to extract results and conclusions. Lists of Minimal pair words are available in a database accessible to Japañol.

## 3. Using the tool

Most of current CAPT systems offer isolated pronunciation or discrimination activities as part of the training exercises . Very few combine these different modes as we do in our learning application. In Japañol we follow a learning methodology based on the Theory, Exposure, Discrimination, Pronunciation and Mixed modes.
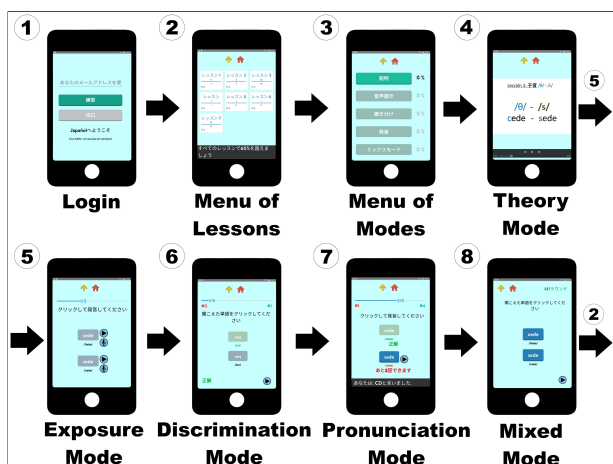
Figure 2: *Standard flow to complete a lesson in Japañol.*

The activities are organized as a sequence of lessons, each devoted to a specific segmental pronunciation difficulty associated to a minimal pair. In each lesson, a brief and clear explanation about the problem and valid pronunciations is provided, in the form of audiovisual material. Then, an exposure mode is entered, in which the user can listen to reference realizations of each valid utterance. After exposure, a discrimination mode is faced in which a sequence of 10 pairs of distinct words (part of a minimal pair) is presented while the user is required to selected which one corresponds best with the listened utterance (generated using the TTS). A minimum number of 6 success hits has to be obtained in order to proceed to the next step. If not, the user is suggested to return to exposure mode again before facing a new discrimination challenge for the same lesson. Once the minimum required number of right answers has been given, or after a number of 5 tries to avoid discouraging the user, a pronunciation mode is entered. In this mode, the user has to say, in sequence, 10 different words selected from the list associated to the minimal pair in the lesson. The recorded speech is submitted to Google Speech ASR and is accepted as valid only when the first item of the N-best list provided by the ASR matches the target word. A minimum number of 6 correct pronunciations is required to pass. If the attempt fails, the user is recommended to return to exposure mode before attempting again.

Finally, a mixed mode activity is required for each lesson. In this mode, a sequence of 10 random discrimination and production tasks are presented and a 60% success is again required to proceed. This mode resembles the added difficulty found to switch from discrimination to production in a normal conversation. A diagram showing the sequence of activities to complete a lesson is shown in Figure 2.

## 4. Activities in the demonstration

The demonstration will consist on an interactive session showing all different modes in the client application (see 3). People will be able to ask for help during the presentation. At the beginning, all attending people can download the application with a given URL or taking a photo of a QR picture. Once downloaded, the demonstration begins logging into the application before entering the menu of lessons.

## 5. Acknowledgements

## 6. References

[1] R. I. Thomson and T. M. Derwing, "The effectiveness of L2 pronunciation instruction: A narrative review," *Applied Linguistics*, vol. 36, no. 3, p. 326, 2014.

[2] W. Li and D. Mollá-Aliod, "Computer processing of oriental languages. language technology for the knowledge-based economy," *Lecture Notes in Computer Science*, vol. 5459, 2009.

[3] Japañol Project. (2017) Eca-simm japañol project. [Online]. Available: https://eca-simm.uva.es/es/proyectos/capt/japanol/

[4] 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!" *Interspeech*, pp. 1178–1179, September 2016.

[5] C. Tejedor-García, D. Escudero-Mancebo, E. Cámara-Arenas, C. González-Ferreras, and V. Cardeñoso-Payo, "Improving L2 production with a gamified computer-assisted pronunciation training tool, TipTopTalk!" *IberSPEECH 2016: IX Jornadas en Tecnologías del Habla and the V Iberian SLTech Workshop events*, pp. 177–186, 2016.

[6] ——, "TipTopTalk! mobile application for speech training using minimal pairs and gamification," *IberSPEECH 2016: IX Jornadas en Tecnologías del Habla and IV Iberian SLTech Workshop events*, pp. 425–432, 2016.

[7] C. Tejedor-García, V. Cardeñoso-Payo, E. Cámara-Arenas, C. González-Ferreras, and D. Escudero-Mancebo, "Playing around minimal pairs to improve pronunciation training," *IFCASL 2015*, 2015.

[8] 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*, no. 1, pp. 355–363, 2018.

[9] 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," *SLaTE*, pp. 125–130, 2015.

[10] 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," *Abstracts of New Sounds Aarhus, 8th International Conference on Second Language Speech*, p. 160, 2016.

[11] 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," *SLaTE*, 2017.

[12] C. Tejedor-García, D. Escudero-Mancebo, E. Cámara-Arenas, C. González-Ferreras, and V. Cardeñoso-Payo, "Improving pronunciation of spanish as a foreign language for l1 japanese speakers with japañol capt tool," *IberSPEECH 2018: X Jornadas en Tecnologías del Habla and the V Iberian SLTech Workshop*, pp. iii–lll, 2018.