

# Diseño y Evaluación de una Aplicación Móvil para el Entrenamiento de la Pronunciación de Lengua Extranjera Basada en Técnicas de Pares Mínimos

## *Design and Evaluation of a Mobile Application for Second Language Pronunciation Training based on Minimal Pairs*

Cristian Tejedor-García  
ECA-SIMM

Departamento de Informática  
Universidad de Valladolid

cristian@infor.uva.es

**Resumen:** Una correcta pronunciación es clave para lograr un nivel de comunicación adecuado en cualquier lengua extranjera. A pesar de ello, los sistemas de aprendizaje tradicionales suelen enfatizar el desarrollo de las competencias lingüísticas gramaticales o léxicas. La calidad de la tecnología del habla (síntesis y reconocimiento de voz) ha mejorado notablemente en los últimos años, surgiendo cada vez más aplicaciones de enseñanza de lengua extranjera que la incorporan. Los juegos serios constituyen un recurso con gran potencial educativo y las versiones sociales animan a una participación competitiva y colaborativa de individuos y grupos. Este proyecto de tesis se centra en el análisis, desarrollo y evaluación de un sistema para la mejora de la pronunciación de lengua extranjera. Se propone un juego serio implementado sobre plataformas móviles y una metodología de aprendizaje que utiliza las tecnologías del habla y motiva al usuario a seguir practicando a lo largo del tiempo, de manera controlada.

**Palabras clave:** Entrenamiento de la pronunciación asistida por ordenador, pares mínimos, aprendizaje de segunda lengua, aprendizaje de idiomas, entrenamiento de la pronunciación, gamificación, juegos serios

**Abstract:** Correct pronunciation is crucial to grasp an adequate communication ability in a foreign language. Nevertheless, traditional foreign language learning systems usually focus on the development of linguistic competencies related to grammar or lexicon. The quality of spoken language technologies (speech synthesis and recognition) has noticeably improved and, as a consequence, an increasing number of second language learning tools have included it. Serious games show great potential as educational resources. Their social versions encourage competitive and collaborative learning scenarios for individuals and groups. The proposal of this thesis project focuses on developing a system for improving foreign language pronunciation. A serious game to be run on a mobile platform is proposed, and a learning method which makes use of speech technology and is designed to guide and motivate users to a continuous practice over time, in a controlled way.

**Keywords:** Computer-assisted pronunciation training, minimal pairs, second learning acquisition, language learning, pronunciation training, gamification, serious games

### 1 *Justificación de la investigación propuesta*

El debate sobre la contribución de las TIC al aprendizaje de lengua extranjera (L2) frente a los métodos tradicionales (Thomson y Derwing, 2014) ha aumentado a lo largo de la última década. Esto ha motivado la aparición de nuevas propuestas basadas en el uso

de aplicaciones para dispositivos móviles inteligentes (Kukulka-Hulme, 2012; Escudero-Mancebo y Carranza, 2015).

En comparación con la atención prestada al aprendizaje asistido por ordenador de los aspectos léxicos (Yoon, 2008) y gramaticales (Levy, 2009), hay pocos estudios que aborden el estudio experimental de la contribu-

ción del uso de las tecnologías del habla al entrenamiento y mejora de la pronunciación (CAPT, en inglés), si bien se reconoce que el uso de reconocedores y sintetizadores de voz puede ayudar a la enseñanza y mejora de la pronunciación en L2 (Hincks, 2005; Neri et al., 2008; Luo, 2016).

El éxito de las tecnologías del habla en el aprendizaje de L2 está directamente relacionado con la elección de la metodología de entrenamiento correcta, que deberá atender adecuadamente los aspectos de realimentación automática instantánea y de un adecuado diseño de las actividades, para optimizar el tiempo de uso de las herramientas CAPT (Neri, Cucchiarini, y Strik, 2008). En este trabajo optamos por la metodología basada en el uso de pares mínimos.

Un par mínimo (Celce-Murcia, Brinton, y Goodwin, 1996) se compone de dos palabras (normalmente monosilábicas), que difieren en un único sonido (e.g. «cat»-«cut» o «bin»-«pin» en inglés). Aunque originalmente se emplearon en tareas de extracción de catálogos fonémicos de lenguas relativamente desconocidas, es una técnica que se aplica en la enseñanza de la pronunciación de segundas lenguas (Baker, 2006). Los pares mínimos facilitan la realización de actividades de exposición a los sonidos de los fonemas, para facilitar la familiarización con los mismos, tareas de percepción entre reproducciones sonoras de pares y, finalmente, actividades de producción de locuciones asociadas a cada uno de ellos (Cámara-Arenas, 2013). La discriminación entre palabras de pares mínimos constituye un desafío para los estudiantes de lenguas extranjeras y para las tecnologías de reconocimiento automático de habla actuales.

La «gamificación» es el uso de estrategias (pensamientos y mecánicas) de juego en entornos ajenos al juego (Deterding et al., 2011). Los videojuegos se están incorporando a los recursos de aprendizaje en los últimos años (Wiggins, 2018), por las posibilidades de cooperación, competición y trabajo individual que aportan y por el potencial incremento de la motivación y participación de los alumnos fuera del aula (Shaffer et al., 2005).

Si bien la inclusión de los elementos competitivos en cualquier sistema pedagógico debe realizarse con precaución, son varios los estudios que indican que la competitividad en el contexto del aprendizaje basado en juegos facilita la consecución de objetivos de apren-

dizaje (Chen, Law, y Chen, 2018) y fomenta la cooperación como elemento articulador del trabajo en clase (Brown y Lee, 2015).

## 2 *Antecedentes y trabajo relacionado*

En la revisión del estado de la cuestión acerca de la enseñanza de pronunciación de lengua extranjera de Derwing y Munro (2005), solo aparecen 26 experimentos empíricos reproducibles hasta 2005, y hasta 2014 solo el 39 % de los 49 experimentos revisados utilizaban tecnologías CAPT.

Los sistemas CAPT iniciales se limitaban a ejercicios de escucha, repetición y elección de sonidos mediante un ordenador (Wang, 2002). Progresivamente, se van incorporando sistemas de reconocimiento automático de voz (ASR) y de síntesis de habla (TTS) como consecuencia de las potenciales ventajas que puede aportar el uso de esta tecnología: evaluación dinámica, realimentación individualizada automática e inmediata, práctica más intensiva, contexto libre de ansiedad y oportunidades para la corrección de errores (Neri et al., 2002; Neri, Cucchiarini, y Strik, 2008).

Para evaluar la posible mejora de la pronunciación utilizando CAPT se pueden seguir dos métodos: asignar valores subjetivos de evaluadores humanos (Mak et al., 2003; Neri et al., 2008), o puntuar de manera objetiva cada pronunciación partiendo de parámetros estadísticos (como el «Goodness of Pronunciation» (Neri, Cucchiarini, y Strik, 2008) o puntuaciones de confianza de reconocedor de voz (Hincks, 2005)). En ambos casos se suele realizar un pre-test y dos post-test, uno tras el entrenamiento y otro retardado. Además, se recurre a estudios de correlación entre los resultados de evaluación subjetiva y objetiva.

Existen experimentos CAPT con realimentación muy simple (y en muchos casos insuficiente) como por ejemplo, reducir el resultado a correcto, incorrecto o puntuaciones de la pronunciación (Neri et al., 2008; Jayakumar et al., 2016). Por otra parte, existen sistemas más sofisticados, tanto gratuitos (*Google ASR*) como privativos (*Nuance*), que se pueden personalizar para presentar una realimentación adaptada a cada usuario, incluyendo el texto del habla reconocido, consejos visuales o textuales sobre cómo conseguir pronunciar correctamente o realizar ejercicios que lo faciliten (Tomokiyo, Wang, y Eskenazi, 2000; Neri, Cucchiarini, y Strik, 2008).

En el ámbito de la «gamificación», apenas se encuentran aún experimentos en el aula de sobre la importancia de la interacción social en el aprendizaje de segundas lenguas, aunque aparecen ya estrategias y consejos para llevarlos a cabo (Flores, 2015), entre las que se considera la competitividad. La tradición establece que a través del juego el educador puede incorporar distensión y variedad a las actividades formativas, lo que promueve que algunos autores concedan un mayor protagonismo a las técnicas lúdico-pedagógicas (Hadfield, 2000) y a la presencia de elementos competitivos que no comprometan la diversión como factor principal (Greenall, 1984).

### 3 Descripción de la investigación propuesta

El objetivo central de la tesis es analizar, comparar y proponer un sistema de guía y evaluación automática de la pronunciación orientada a estudiantes de lengua extranjera, utilizando síntesis y reconocimiento automático del habla para la consecución y mejora de los ejercicios de pronunciación, facilitando así el aprendizaje autónomo del estudiante en un entorno privado, sin situaciones de ansiedad, a su propio ritmo y con la posibilidad de recibir realimentación instantánea en cualquier momento y lugar. Este objetivo general conlleva la consecución de los siguientes objetivos particulares:

- **O1:** Estudio y selección de actividades y metodologías para la mejora de pronunciación de lengua extranjera.
- **O2:** Estudio y selección de los criterios y parámetros de caracterización del habla más adecuados para la clasificación de nivel de pronunciación de las personas.
- **O3:** Análisis, desarrollo y evaluación de un método de obtención de un conjunto de pares mínimos específicos de problemas de pronunciación extranjera según la L1, supervisado por experto.
- **O4:** Análisis y diseño de algoritmos específicos de reconocimiento de voz para la identificación de errores de pronunciación asociados a parámetros clave de caracterización de nivel.
- **O5:** Análisis, desarrollo y evaluación de un sistema que proporcione realimentación individualizada al usuario para su mejora de la pronunciación extranjera.

### 4 Metodología y experimentos propuestos

Tras la revisión pormenorizada del estado de la cuestión, se ha concluido que la selección de la metodología de entrenamiento de la pronunciación es un elemento clave. En este sentido, la decisión de comenzar por una estrategia basada en el uso de pares mínimos está demostrando ser fructífera. No obstante, desde los primeros pasos ha resultado evidente que la elección de los pares que el participante debe reconocer y pronunciar es un aspecto clave y que dicha selección de pares debe estar adaptada, necesariamente, a las características de su L1. Habida cuenta de que se están desarrollando aplicaciones y experimentos para diversos pares L1-L2, se ha hecho necesario también el diseño de un algoritmo para la preparación de candidatos de pares mínimos en el idioma L2 objetivo, que luego puedan ser revisados por un experto (Tejedor-García y Escudero-Mancebo, 2018).

Como experimento inicial, se ha diseñado una versión de aplicación para uso individual, de cara a valorar la utilidad de la tecnología del habla (ASR y TTS) para clasificar a los usuarios según su nivel de competencia en pronunciación (Escudero-Mancebo et al., 2015).

En una segunda etapa, se han incorporado elementos de «gamificación» a la aplicación móvil para poder valorar tanto aspectos de facilidad de uso como, sobre todo, el nivel de mejora de la pronunciación alcanzable por medio de esta aproximación y usando pares mínimos como metodología de aprendizaje (Tejedor-García et al., 2015; Tejedor-García et al., 2016; Tejedor-García et al., 2016). Los resultados mostraron claramente que el uso de esta aplicación favorece una mejora de la pronunciación en un entorno de juego que resulta fácil de usar y atractivo.

En una tercera etapa, se están incorporando aspectos competitivos en el juego, para estudiar su influencia sobre la motivación del usuario y la mejora de la pronunciación. En esta versión, la aplicación se plantea como un juego de carácter competitivo en el que los usuarios pueden obtener logros y puntuaciones y lanzar retos de competición a otros usuarios como parte de la dinámica de uso de la aplicación de mejora de la pronunciación. Un análisis preliminar de estos resultados están pendientes de revisión.

En paralelo con las dos etapas anteriores,

se han desarrollado también variantes de la aplicación orientadas a un uso pautado de la misma, definido por expertos en enseñanza de la lengua extranjera L2 en cada caso. El objetivo de estos experimentos es proporcionar datos y conclusiones sobre la forma de mejorar la realimentación a los usuarios en caso de fallo. Los resultados de estos experimentos en configuración guiada y controlada están también en proceso de revisión.

Como última etapa, se plantea un estudio sistemático de las características del habla adquirido durante los experimentos realizados para poder determinar el conjunto de rasgos clave a la hora de poner en correlación los fallos de pronunciación con las valoraciones de expertos y, por otro lado, para facilitar la extracción a partir de ello de sistemas de clasificación automática que puedan guiar la formulación de recomendaciones personalizadas sobre modo y punto de articulación.

## 5 Cuestiones de investigación

Además de la discusión sobre metodología y resultados obtenidos, son varias las cuestiones de investigación que se plantean, entre las que destacan las siguientes:

1. ¿Qué impacto puede tener el uso de un reconocedor de voz comercial frente a sistemas personalizados, diseñados explícitamente para la mejora de pronunciación, y a qué nivel debe abordarse la adaptación (usuario, lengua L2)?
2. ¿Qué nivel de mejora de pronunciación puede lograrse en un plazo de exposición más largo y qué impacto tiene la metodología empleada?
3. ¿Qué influencia tienen los elementos de juego y la configuración competitiva sobre la motivación, el nivel de aprendizaje, la auto-conciencia de nivel, o el grado de ansiedad o frustración del usuario?

## Agradecimientos

Este trabajo ha sido parcialmente financiado por Ministerio de Economía y Competitividad y Fondos FEDER – clave de proyecto: TIN2014-59852-R Videojuegos Sociales para la Asistencia y Mejora de la Pronunciación de la Lengua Española, por la Universidad de Valladolid (Convocatoria 2015 de Contratos Predoctorales) y por la Consejería de Educación de Junta de Castilla y León en el proyecto (VA050G18).

## Bibliografía

- Baker, A. 2006. *Ship Or Sheep? Student's Book: An Intermediate Pronunciation Course*, volumen 1. Ernst Klett Sprachen.
- Brown, H. D. y H. Lee. 2015. *Teaching by Principles: An Interactive Approach to Language Pedagogy*, volumen 1. Prentice Hall Regents Englewood Cliffs, NJ.
- Celce-Murcia, M., D. Brinton, y J. M. Goodwin. 1996. *Teaching Pronunciation*.
- Chen, C.-H., V. Law, y W.-Y. Chen. 2018. The Effects of Peer Competition-based Science Learning Game On Secondary Students' Performance, Achievement Goals, And Perceived Ability. *Interactive Learning Environments*, 26(2):235–244.
- Cámara-Arenas, E. 2013. *Native Cardinality: on Teaching American English Vowels to Spanish Students*. Historia y sociedad. Ediciones Universidad de Valladolid.
- Derwing, T. M. y M. J. Munro. 2005. Second Language Accent and Pronunciation Teaching: A Research-Based Approach. *TESOL Quarterly*, 39(3):379–397.
- Deterding, S., D. Dixon, R. Khaled, y L. Nacke. 2011. From Game Design Elements to Gamefulness: Defining Gamification. En *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, páginas 9–15. ACM.
- Escudero-Mancebo, D. y M. Carranza. 2015. Nuevas Propuestas Tecnológicas para la Práctica y Evaluación de la Pronunciación del Español como Lengua Extranjera. En *Congreso Internacional de la Asociación Europea de Profesores de Español*, páginas 218–227, Burgos, Spain. July 20–24.
- Escudero-Mancebo, D., E. Cámara-Arenas, C. Tejedor-García, C. González-Ferreras, y V. Cardenoso-Payo. 2015. Implementation and Test of a Serious Game Based on Minimal Pairs for Pronunciation Training. *SLaTE*, páginas 125–130.
- Flores, J. F. F. 2015. Using Gamification To Enhance Second Language Learning. *Digital Education Review*, 1(27):32–54.
- Greenall, S. 1984. *Language Games and Activities*. Hulton educational publications Ltd, Britain.

- Hadfield, J. 2000. *Beginners' Communication Games*. Longman.
- Hincks, R. 2005. *Computer Support for Learners of Spoken English*. Ph.D. tesis, KTH.
- Jayakumar, A., M. Raghunath, M. Sakthipriya, S. Akhila, A. Sadanandan, y P. Nedungadi. 2016. Enhancing Speech Recognition in Developing Language Learning Systems for Low Cost Androids. En *ICTICT*, páginas 80–84. IEEE.
- Kukulska-Hulme, A. 2012. *Mobile-Assisted Language Learning*. Blackwell Publishing Ltd.
- Levy, M. 2009. Technologies in Use for Second Language Learning. *The Modern Language Journal*, 93(s1):769–782.
- Luo, B. 2016. Evaluating a Computer-Assisted Pronunciation Training (CAPT) Technique for Efficient Classroom Instruction. *Computer Assisted Language Learning*, 29(3):451–476.
- Mak, B., M. Siu, M. Ng, Y.-C. Tam, Y.-C. Chan, K.-W. Chan, y Leung. 2003. PLASER: Pronunciation Learning via Automatic Speech Recognition. En *Proc. HLT-NAACL*, páginas 23–29. Association for Computational Linguistics.
- Neri, A., C. Cucchiari, y H. Strik. 2008. The Effectiveness of Computer-Based Speech Corrective Feedback for Improving Segmental Quality in L2-Dutch. *RECALL*, 20(02):225–243.
- Neri, A., C. Cucchiari, H. Strik, y L. Boves. 2002. The Pedagogy-Technology Interface in Computer-Assisted Pronunciation Training. *Computer Assisted Language Learning*, 15(5):441–467.
- Neri, A., O. Mich, M. Gerosa, y D. Giuliani. 2008. The Effectiveness of Computer Assisted Pronunciation Training for Foreign Language Learning by Children. *Computer Assisted Language Learning*, 21(5):393–408.
- Shaffer, D. W., K. R. Squire, R. Halverson, y J. P. Gee. 2005. Video Games and the Future of Learning. *Phi Delta Kappan*, 87(2):105–111.
- Tejedor-García, C., V. Cardeñoso-Payo, E. Cámara-Arenas, C. González-Ferreras, y D. Escudero-Mancebo. 2015. Playing Around Minimal Pairs to Improve Pronunciation Training. *IFCASL 2015*.
- Tejedor-García, C., V. Cardeñoso-Payo, E. Cámara-Arenas, C. González-Ferreras, y D. Escudero-Mancebo. 2016. Measuring Pronunciation Improvement in Users of CAPT tool TipTopTalk! *Interspeech*, páginas 1178–1179, Septiembre.
- Tejedor-García, C. y D. Escudero-Mancebo. 2018. 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*, 1(1):355–363.
- Tejedor-García, C., D. Escudero-Mancebo, E. Cámara-Arenas, C. González-Ferreras, y V. Cardeñoso-Payo. 2016. Improving L2 Production with a Gamified Computer-Assisted Pronunciation Training Tool, TipTopTalk! *IberSPEECH 2016*, páginas 177–186.
- Thomson, R. I. y T. M. Derwing. 2014. The Effectiveness of L2 Pronunciation Instruction: A Narrative Review. *Applied Linguistics*, 36(3):326.
- Tomokiyo, L. M., L. Wang, y M. Eskenazi. 2000. An Empirical Study of the Effectiveness of Speech-Recognition-Based Pronunciation Training. En *Proceedings of the 6th ICSLP volume*, páginas 677–680.
- Wang, X. 2002. *Training Mandarin And Cantonese Speakers To Identify English Vowel Contrasts: Long-term Retention And Effects On Production*. Ph.D. tesis, Simon Fraser University.
- Wiggins, B. E. 2018. An Overview and Study on the Use of Games, Simulations, and Gamification in Higher Education. En *Gamification in Education: Breakthroughs in Research and Practice*. IGI Global, páginas 191–204.
- Yoon, H. 2008. More than a Linguistic Reference: The Influence of Corpus Technology on L2 Academic Writing. *Language Learning & Technology*, 12(2):31–48.