

## USO DE PARES MÍNIMOS EN HERRAMIENTAS PARA LA PRÁCTICA DE LA PRONUNCIACIÓN DEL ESPAÑOL COMO LENGUA EXTRANJERA

Cristian Tejedor-García y David Escudero-Mancebo  
*Departamento de Informática, Universidad de Valladolid (España)*

Un par mínimo es un conjunto de dos palabras que difieren en sólo uno de los fonemas que constituyen su producción oral, cambiando por completo su significado. Existen programas informáticos que emplean pares mínimos para el entrenamiento de la pronunciación de lengua extranjera, principalmente para el inglés. En este artículo se presenta una herramienta que utiliza voz sintética y un sistema de reconocimiento automática del habla para, en combinación con un ciclo de exposición, discriminación y producción de pares mínimos, entrenar la pronunciación de idiomas extranjeros. Para adaptar estas herramientas al español es necesario, primero disponer de una lista de pares mínimos, y segundo, elegir los pares mínimos que pueden ser interesantes para determinados tipos de hablantes de español como lengua extranjera, en función siempre de su idioma nativo. En este trabajo se presenta una estrategia de selección de pares mínimos basada en métodos automáticos. Un algoritmo encuentra en un diccionario y en una serie de textos el conjunto total de pares mínimos del español. Un nuevo algoritmo, filtra y elige los pares mínimos más adecuados en función del par de fonemas concreto con el que se quiera trabajar. Finalmente, el profesor selecciona las actividades más apropiadas de entre aquellas propuestas por el programa informático. El resultado final es un programa informático de fácil manejo que permite configurar actividades de pronunciación del español para estudiantes ELE de diversas lenguas maternas. **Palabras clave:** tecnología del habla, pares mínimos, pronunciación L2, reconocimiento automático de voz, juego serio.

## INTRODUCCIÓN

A lo largo de la última década, la hipótesis de que las Tecnologías de la Información y la Comunicación (TICs) y la tecnología del habla en particular están preparadas para contribuir a la enseñanza de segunda lengua extranjera (L2) ha alimentado mucho debate y llevado a una serie de interesantes propuestas prácticas (Ehsani 1998, Kukulska-Hulme 2012, Thomson 2014, Escudero-Mancebo 2015). Mientras que la adquisición asistida por ordenador de aspectos como la gramática (Yoon 2008, Levy 2009, Sauro 2009) y vocabulario (Al-Seghayer 2005, Levy 2009) L2 ha recibido una constante atención académica, ha habido relativamente pocos intentos de determinar empíricamente si la tecnología, y la tecnología del habla en particular, con herramientas Computer-Assisted Pronunciation Training (CAPT) como los sintetizadores y reconocedores de voz, pueden ayudar eficazmente a entrenar, enseñar y mejorar la pronunciación L2 (Hincks 2005, Neri 2008a, Neri 2008b, Luo 2016). Desde un punto de vista pedagógico, el uso de pares mínimos permite a los estudiantes darse cuenta de los riesgos potenciales de cambiar un único fonema al pronunciar dos palabras tan similares (Celce-Murcia 1996). Distinguir entre las palabras que componen el par mínimo constituye, a priori, una tarea difícil para el sistema de reconocimiento automático del habla, ya que la distancia fonética entre ambas es muy pequeña. Esto es, las palabras del par mínimo son fácilmente confundibles si la producción no es lo suficientemente clara y precisa.

Primero se detallan algunas herramientas software para la mejora de la pronunciación de idiomas extranjeros que utilizan pares mínimos, junto a sus actividades y metodologías más características. Después se describe una herramienta propia adaptada al castellano, con unas características especiales de juego serio. Posteriormente se explica un algoritmo automático de obtención de pares mínimos de palabras procedentes de textos en castellano. Por último se presentan la discusión y conclusiones del artículo.

## HERRAMIENTAS SOFTWARE

En los últimos años han aparecido nuevas herramientas informáticas dedicada a la enseñanza de segunda lengua, en concreto a la pronunciación del español (Escudero-Mancebo 2015). En concreto, existen muy pocas herramientas que se basen en el uso de pares mínimos para el castellano, siendo la mayoría son para el idioma inglés. Por ejemplo, el programa para ordenador Minimal Pairs Tutor de la empresa Rose Medical (ver figura 1), muestra en su interfaz los pares mínimos con una imagen simplificada y representativa de su significado. Clasifica los pares mínimos en función del fonema cambiante, y los muestra en diferentes pantallas. La dinámica de esta herramienta es la siguiente: se debe pulsar en la imagen de cada palabra y elegir la opción de grabarse para obtener un reconocimiento en forma de texto y una puntuación de inteligibilidad. También se puede escuchar la pronunciación nativa de las palabras.

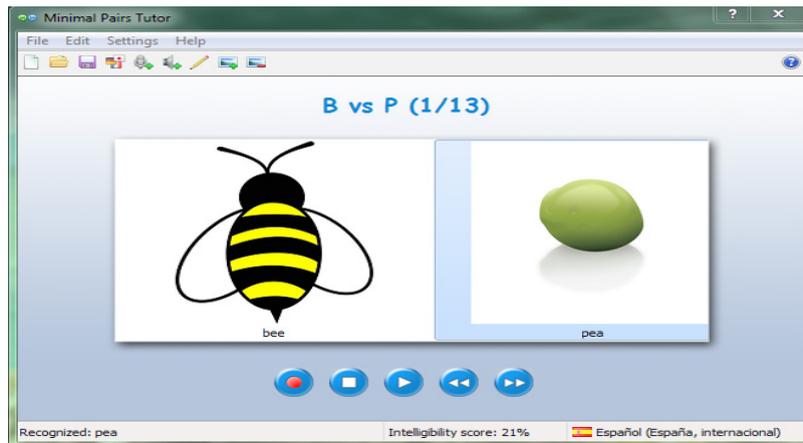


Figura 1: Interfaz de ejemplo de pronunciación de par mínimo en inglés en la herramienta *Minimal Pairs Tutor*.

Otra herramienta que utiliza pares mínimos es *MinimalPairs* (ver figura 2), de la empresa Mobi-Learning Inc, en este caso para dispositivos móviles Android. Posee actividades aisladas de percepción auditiva, discriminación perceptiva y producción. En las del primer tipo, la aplicación muestra en una gran matriz todas las palabras asociadas a los fonemas del par mínimo. El usuario pulsa y escucha cada una de ellas como desee. Por otra parte, en las actividades de discriminación perceptiva, se muestran las imágenes de las palabras (sin el texto) a la vez que se escucha una de ellas al azar. El usuario deberá escoger cuál cree que ha sido la que ha escuchado. Por último, en las actividades de producción el usuario pulsa sobre la imagen de la palabra y la pronuncia. El sistema acepta o rechaza dicha pronunciación sin ofrecer más realimentación que un correcto o incorrecto.

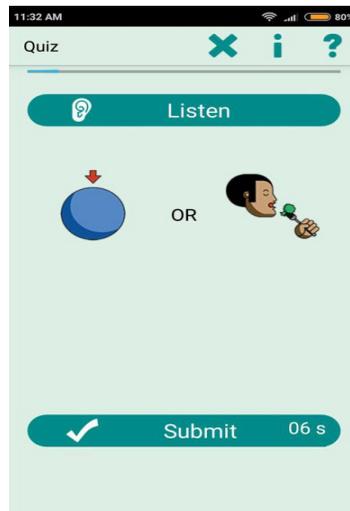


Figura 2: Interfaz de una actividad de discriminación auditiva de un par mínimo en la herramienta *MinimalPairs*.

TipTopTalk! es una aplicación móvil que consiste en un juego serio para la mejora de la pronunciación mediante pares mínimos (Tejedor-García 2016a, b, c). Incluye, entre otros idiomas, la práctica del español como lengua extranjera. Esta herramienta se basa parcialmente en el ciclo exposición-discriminación-pronunciación del Método de la Cardinalidad Nativa (Cámara-Arenas 2012). La peculiaridad de esta aplicación es que recomienda ciertos ejercicios a los usuarios en función de sus resultados (no de manera aislada), adaptándose a ellos.

En las actividades de exposición de la herramienta TipTopTalk! (ver la primera imagen de la figura 3) los usuarios son ayudados a familiarizarse con los sonidos de los fonemas mediante secuencias de pares mínimos que aparecen de manera aleatoria. Las palabras son sintetizadas un máximo de cinco veces, cada vez más despacio, intercalando las palabras. Tras ello, los usuarios deben grabarse pronunciando la palabra al menos una vez para comparar ambos sonidos. El siguiente modo de entrenamiento es el de discriminación o percepción auditiva (ver la segunda imagen de la figura 3), en el que los usuarios comprueban su habilidad de percepción entre el fonema cambiante de las palabras de los pares mínimos. En primer lugar escuchan la síntesis de voz de una de las palabras del par mínimo, y deben discernir cuál es. Una vez elegida, automáticamente aparece un nuevo par mínimo. En último lugar, en el modo de producción (ver la tercera imagen de la figura 3), los usuarios deben pronunciar las palabras del par mínimo correctamente. Para ello se graban la voz pulsando sobre la palabra que desean (pueden escuchar una síntesis de las mismas tantas veces como quieran de manera opcional). Una realimentación en tiempo real es mostrada por pantalla, en forma de mensaje. En caso de que no sea correcta dicha pronunciación, se muestra el resultado de la misma y un consejo breve para motivar al usuario. Se permiten hasta cinco fallos por palabra para evitar inferir en la frustración y el agotamiento del estudiante.

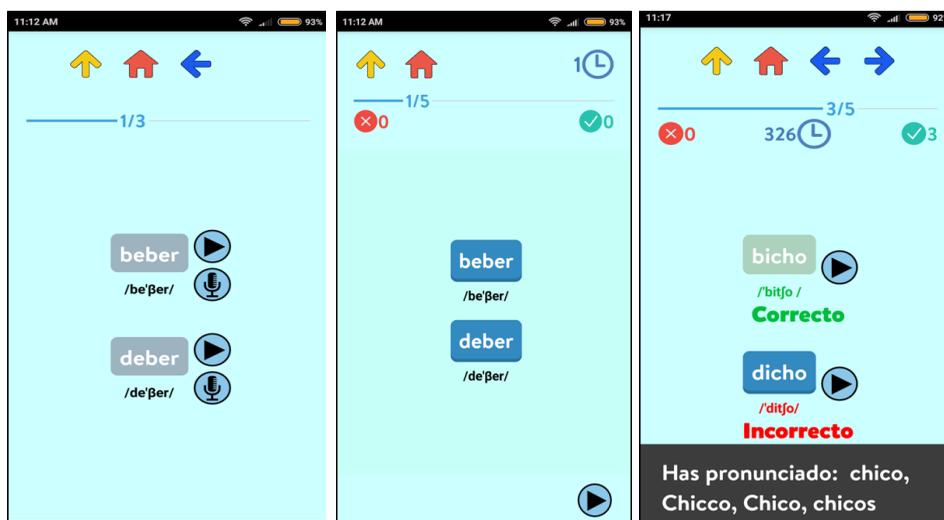


Figura 3: Ciclo de actividades de exposición-discriminación-pronunciación de la herramienta TipTopTalk!

Los elementos de juego incluidos en las actividades con las que se obtienen puntos en TipTopTalk! permiten a los usuarios participar en *rankings* y conseguir trofeos, proporcionando motivación y competitividad, a la vez que se sigue aprendiendo. No obstante, conceder a los estudiantes elegir qué actividades realizar con total flexibilidad puede acarrear problemas como estancamiento en aquellas pruebas más difíciles. Por ello se proponen baterías de actividades guiadas en función de la interactividad y resultados obtenidos.

## OBTENCIÓN AUTOMÁTICA DE PARES MÍNIMOS

Para obtener pares mínimos de manera automática, se ha diseñado un algoritmo recursivo propio que compara los fonemas de las palabras de una serie de textos o diccionarios en español. Para poder aplicar el algoritmo, se deberán dar unos pasos previos y posteriores para poder obtener así el resultado final esperado: listas de pares mínimos indexadas por fonemas problemáticos para el idioma objetivo, y adaptables a la tecnología.

En la figura 4 se muestran dichos pasos en un esquema de estados. El primer paso consiste en escoger el idioma nativo L1 y el extranjero L2 objetivos. Este paso es fundamental, ya que hay que caracterizar el idioma extranjero en función del habla materna. El segundo paso y tercer paso se pueden realizar en paralelo. En el caso del segundo paso, el corpus de textos será elegido previamente en función del idioma y de las características de las palabras con las que se quiera trabajar, desde textos científicos, palabras o frases aisladas, a novelas literarias. Para el caso concreto del castellano, se han utilizado los textos de diversas novelas recientes debido a que contienen un amplio y variado léxico actual (incluyendo algunas obras de Benito Pérez Galdós). En total, los textos utilizados recogen aproximadamente tres millones de palabras. Por otra parte, el tercer paso trata en realizar una revisión de la literatura respecto al estado de la cuestión acerca de los fonemas más complejos del idioma castellano en función del idioma nativo de sus estudiantes.

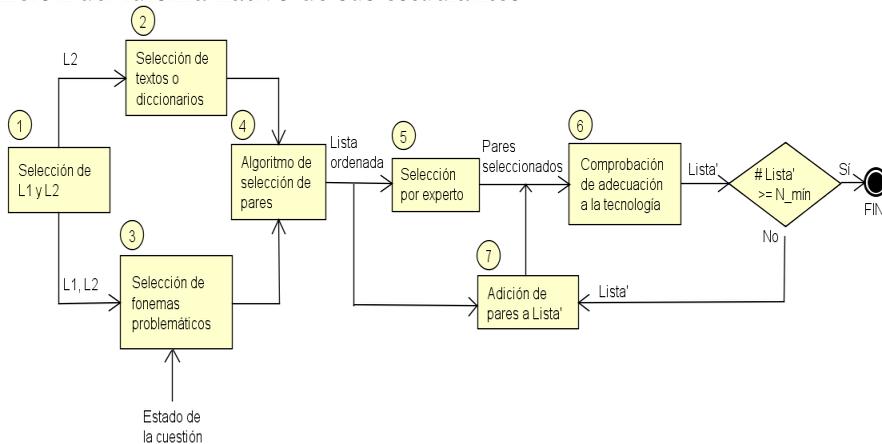


Figura 4: Pasos para la obtención de listas de pares mínimos de palabras susceptibles de ser incluidas en herramientas informáticas de mejora de la pronunciación de lengua extranjera. Cada uno de los rectángulos son estados posibles, siendo el rombo un estado de decisión y el punto negro del extremo derecho el fin del diagrama.

El cuarto paso consiste en aplicar (ejecutar) el algoritmo para la búsqueda de los pares. El algoritmo consiste en la lectura secuencial y elección específica de las palabras obtenidas de un corpus de textos especificados. En concreto, primero se crea un fichero único con las palabras, su representación fonética SAMPA y su frecuencia de aparición. Después se realiza una búsqueda en árbol recursivo, de manera que se van descartando palabras que no pueden formar pares mínimos, mirando el tamaño de la palabra y cada fonema que la forma, pudiendo ser descartada antes de ser recorrida por completo. Se obtiene así otro fichero con la lista de las palabras de los pares mínimos, junto a su representación fonética, ordenados por su frecuencia de aparición y clasificados por fonema cambiante.

En términos informáticos, la complejidad temporal de un algoritmo es el tiempo que consume un algoritmo para resolver un problema. La complejidad del algoritmo en el tiempo es lineal (su ejecución sería aproximadamente segundos o minutos según el número de palabras total):  $O(N \cdot L \cdot L(F))$ , siendo  $N$  el número de palabras de longitud máxima  $L$ , y  $L(F)$  longitud máxima de palabra en función del número de fonemas del idioma. En el caso concreto de este artículo, aproximadamente ha tardado un minuto en tiempo de ejecución. La alternativa sería la denominada en el argot informático “fuerza bruta”, que ha sido descartada en primera instancia, ya que su complejidad en el tiempo es cuadrática (su ejecución sería aproximadamente horas o días según el número de palabras total):  $O(L \cdot N^2)$ , siendo  $N$  el número de palabras de longitud máxima  $L$ .

En el quinto paso los pares son escogidos por un experto (profesor, investigador o similar), utilizando la lista ordenada obtenida del paso anterior. La lista será leída de arriba a abajo (por frecuencia de aparición de las palabras en los textos) y los pares serán escogidos en función de la experiencia del experto, es decir, si ha utilizado dichas palabras con anterioridad en su experiencia profesional, o si tiene algún criterio especial, como podría ser el de elegir palabras según la posición del fonema problemático en las mismas (inicial, intermedia, final, intervocálica o en grupos consonánticos) o por la frecuencia de uso actual en el idioma. Otro factor a tener en cuenta es la adecuación a la tecnología (ver sexto estado de la figura 4). En particular, pueden ser problemáticas las palabras homófonas (aquellas con misma pronunciación, pero distinta escritura y significado). Además, otro tipo de palabras problemático es el de aquellas palabras fuera de contexto o infrecuentes en el idioma, que podrían no haber sido incluidas en el corpus de reconocimiento o síntesis de voz de la herramienta *software* en cuestión (un problema que cada vez es menor porque se dispone de gran cantidad de datos, almacenamiento y recursos informáticos).

Tras el sexto paso se dispone de una lista de pares mínimos (Lista’) que será definitiva si su número de pares mínimos es al menos igual al mínimo número de pares necesarios ( $N_{\min}$ ). En caso de que no lo sea, el experto volverá a añadir pares procedentes de la lista ordenada tras el algoritmo a la Lista’ que le faltan (ver séptimo paso de la figura 4). Una vez completados los pasos anteriores se dispondrá en su totalidad de las listas de pares mínimos de palabras en el *software* de mejora de la pronunciación de lengua extranjera y completamente accesibles para su uso en otros ámbitos como puede ser medios escritos u orales en aulas y/o centros educativos centrado en este ámbito.

## CONCLUSIONES

En este artículo se ha realizado una revisión de herramientas *software* para la práctica de pronunciación del español como lengua extranjera que utilizan pares mínimos de palabras. Como resultado, se han descrito las actividades que se llevan a cabo en dichos programas. Además, se ha presentado un procedimiento automático de obtención de pares mínimos de textos especificados mediante un algoritmo propio. Todo ello desde una perspectiva interdisciplinar en la que participan expertos de distintas áreas como la informática o la lingüística. La tecnología actual es emergente y de rápida implantación, con la posibilidad de adaptarse a otros idiomas en el futuro. No obstante, cada idioma extranjero tendrá unos problemas concretos según el idioma nativo, y no siempre habrá suficientes pares mínimos o no será posible su adaptación a la tecnología (palabras infrecuentes, combinaciones inexistentes de fonemas, palabras aisladas fuera de contexto y palabras homófonas). Además, influye la subjetividad de los expertos, por lo que hay que llevar a cabo experimentación para comprobar y mejorar los resultados de los programas informáticos.

Las herramientas informáticas CAPT que se han descrito tienen en común que utilizan pares mínimos en las actividades para la práctica de la pronunciación de lengua extranjera. La mayoría son específicas del idioma inglés, no obstante, se ha presentado una herramienta propia, TipTopTalk!, que utiliza pares mínimos en castellano en función del idioma nativo del estudiante. Las actividades que se realizan en estas herramientas informáticas para la mejora de la pronunciación son en su mayoría aisladas, y no tienen en cuenta el contexto o nivel del estudiante. Sin embargo, en TipTopTalk! se hace un gran hincapié en sugerir ejercicios en base a los resultados que se vayan obteniendo, individualizando y personalizando al máximo la experiencia de usuario.

Las diversas actividades de exposición, percepción y producción de estas herramientas son exitosas en parte, porque se pueden llevar a cabo en cualquier lugar sin la presencia de otras personas que puedan incomodar a los estudiantes y en cualquier momento. En el caso concreto de TipTopTalk! se incluyen elementos de juego que mejoran la motivación y compromiso de los usuarios a seguir entrenando, a la vez que se divierten y socializan. La interacción social es clave en estos días, por lo que establecer vínculos entre los distintos usuarios ayuda aún más a la utilización de estas herramientas, alargando a su vez su vida útil, ya que si no se adaptan a los constantes cambios de plataformas acabará desapareciendo. No obstante, aunque los sistemas de reconocimiento y síntesis de habla han mejorado mucho con el paso de los años y en la mayoría de casos funcionan correctamente, siempre hay que tener en cuenta sus limitaciones de cara a adecuar las actividades de las herramientas para evitar resultados insatisfactorios y/o desmotivantes.

Por otra parte, se ha presentado un algoritmo de obtención de listas de pares mínimos de manera automática para el castellano en la herramienta TipTopTalk! Aunque el proceso es automático, se debe contar siempre con la colaboración de expertos para decidir qué fonemas escoger, los textos de los cuáles se obtienen las palabras o establecer criterios extra para la elección final de pares. La facilidad de adaptación a cualquier tamaño de textos de búsqueda e idioma resulta útil para su utilización en otros ámbitos pedagógicos y he-

rramientas similares. Sin embargo, no siempre se cuenta con un número suficiente de palabras relacionadas con los pares mínimos, debido a que el propio idioma no dispone de una cantidad mínima de las mismas, o de si no se pueden adaptar a la tecnología.

## BIBLIOGRAFÍA

- Al-Seghayer, K.. “The effect of multimedia annotation modes on L2 vocabulary acquisition”. *Research in technology and second language education: Developments and directions*, 3/133 (2005).
- Cámara-Arenas, E. *Curso de pronunciación de la lengua inglesa para hispano-hablantes, a native cardinality method*. Valladolid: Secretariado de Publicaciones de la Universidad de Valladolid, 2012.
- Celce-Murcia M., Brinton D. M. , y Goodwin J. M. *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge University Press, 1996.
- Ehsani F. y Knodt E. , “Speech technology in computer-aided language learning: Strengths and limitations of a new call paradigm,” *Language Learning & Technology*, 2/ 1 (1998): 45–60.
- Escudero-Mancebo, D. and Carranza, M. “Nuevas propuestas tecnológicas para la práctica y evaluación de la pronunciación del español como lengua extranjera”. *Congreso AEPE* (2015).
- Hincks, R. *Computer support for learners of spoken English*. PhD thesis, KTH (2005).
- Kukulska-Hulme, A. *Mobile-Assisted Language Learning*. Blackwell Publishing Ltd, (2012).
- Levy, M. “Technologies in use for second language learning”. *The Modern Language Journal*, 93/s1 (2009):769–782.
- Neri, A., Cucchiari, C., and Strik, H. . “The effectiveness of computer-based speech corrective feedback for improving segmental quality in L2-Dutch”. *ReCALL*, 20/02 (2008a):225–243.
- Neri, A., Mich, O., Gerosa, M., and Giuliani, D. “The effectiveness of computer assisted pronunciation training for foreign language learning by children”. *Computer Assisted Language*

*Learning*, 21/5 (2008b): 393–408.

Sauro, S. “Computer-mediated corrective feedback and the development of L2 grammar”. *Language Learning & Technology*, 13/1 (2009): 96–120.

Tejedor-García, C., Cardenoso-Payo, V., Cámara-Arenas, E., González-Ferreas, C., and Escudero-Mancebo, D. “Measuring pronunciation improvement in users of CAPT tool TipTopTalk!” *Interspeech 2016*, (2016a): 1178–1179.

Tejedor-García, C., Escudero-Mancebo, D., Cámara-Arenas, E., González-Ferreas, C., y Cardenoso-Payo, V.. “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*, (2016b): 177–186.

Tejedor-García, C., Escudero-Mancebo, D., Cámara-Arenas, E., González-Ferreas, C., and Cardenoso-Payo, V. “TipTopTalk! mobile application for speech training using minimal pairs and gamification”. *IberSPEECH 2016: IX Jornadas en Tecnologías del Habla and the V Iberian SLTech Workshop events*, (2016c): 425–432.

Thomson, R. I. and Derwing, T. M. “The effectiveness of L2 pronunciation instruction: A narrative review”. *Applied Linguistics*, 36/3 (2014): 326.

Yoon, H. “More than a linguistic reference: The influence of corpus technology on L2 academic writing”. *Language Learning & Technology*, 12/2 (2008):31–48.