

Entrenamiento Fonético por Ordenador para la Adquisición de Segundas Lenguas: ¿Son Efectivos estos Sistemas de Aprendizaje?

Daniel Calvo Carmona
Universidad de Sevilla (España)

Laura César Garrido
Universidad de Salamanca (España)

Entrenamiento Fonético por Ordenador para la Adquisición de Segundas Lenguas: ¿Son Efectivos estos Sistemas de Aprendizaje?

Computer-assisted Phonetic Training: Are They Effective?

Daniel Calvo Carmona

Universidad de Sevilla (España)

dcalvo@us.es

Laura César Garrido

Universidad de Salamanca (España)

Laucesgar@usal.es

Fecha de recepción: 27 de octubre de 2017

Fecha de aceptación: 25 de septiembre de 2018

Resumen

El presente estudio revisa una serie de estudios sobre la efectividad del aprendizaje del componente fonético de las segundas lenguas mediante el uso de sistemas asistidos por ordenador, atendiendo a los criterios de la escala de calidad metodológica MQS (Sanduvete, 2008, p. 285). Se realiza una búsqueda exhaustiva en las bases de datos más importantes para obtener un total de 66 estudios e identificar los problemas metodológicos más importantes. Los resultados de la revisión sugieren que la mayoría de los estudios sobre entrenamiento fonético poseen una calidad de media, especialmente debido a la falta de seguimiento, la falta de métodos de enmascaramiento, y que puedan existir variables extrañas que no se han tenido en cuenta. A partir de estos datos, se incluye una propuesta de mejora en forma de una serie de recomendaciones para mejorar la calidad de estudios posteriores.

Palabras claves: Entrenamiento fonético; Segundas lenguas; Calidad metodológica MQS; Sistemas de aprendizaje por ordenador; Sanduvete

Abstract

The present study reviews research on the effectiveness of the acquisition of the phonetics of a second language through the use of computer assisted systems, following the assessment criteria of the methodological quality scale MQS developed by Sanduvete (2008, p. 285). A comprehensive search is conducted in the most important databases to obtain a total of 66 studies and identify potential methodological problems. Results suggest that most studies on phonetic training have average quality, especially due to lack of follow-up, absence of masking methods, and the failure to account for possible extraneous variables. A proposal is included in the form of a series of recommendations to improve study quality in future studies.

Keywords: Phonetic Training; Second Languages; Methodological Quality Scale (MQS); Computer learning systems; Sanduvete.

Para citar este artículo: Calvo Carmona, D. y César Garrido, L. (2019). Entrenamiento Fonético por Ordenador para la Adquisición de Segundas Lenguas: ¿Son Efectivos estos Sistemas de Aprendizaje?. *Revista de humanidades*, n. 36, pp. 181-210. ISBN 1130-5029 (ISSN-e 2340-8995).

Sumario: 1. Introducción. 2. Objetivos. 3. Metodología. 4. Resultados. 5. Modelo de Entrenamiento Metodológico. 6. Conclusiones y discusión. 7. Bibliografía. 8. Anexos.

1. INTRODUCCIÓN

De todos los aspectos que caracterizan el dominio de una segunda lengua, la fonética de la lengua extranjera es uno de los componentes más difíciles de adquirir, en parte por verse fuertemente influido por la lengua materna, que a menudo difiere de la lengua meta en numerosas características. Asimismo, otro de los problemas más grandes en esta adquisición es la escasez y falta de acceso a muestras de la lengua meta que no se encuentren degradadas, tanto por la escasez de hablantes nativos como por la falta de un nivel adecuado de pronunciación en los profesores de la lengua extranjera. Para mitigar este problema, se han desarrollado diversos sistemas de entrenamiento de la fonética asistidos por ordenador, que prometen mejorar la pronunciación de los estudiantes de una segunda lengua supliendo los problemas anteriormente mencionados. Se trata de programas que ofrecen una flexibilidad horaria y de lugar muy amplio, ya que son accesibles a cualquier hora durante un período de tiempo indefinido pudiéndose ajustar a las necesidades de cada estudiante.

La característica principal de estos programas de entrenamiento es la diversidad de tipos de entrenamiento del que se dispone (Iverson et al, 2012, p. 4). El sistema de entrenamiento más común suele ser el de entrenamiento perceptivo (Lively, et al., 1993, p. 2; Nishi, & Kewley-Port, p. 4 (2007); Iverson & Evans, p. 1 2009): se presentan los sonidos de la lengua extranjera en tareas de identificación o de

discriminación con sonidos difíciles de percibir, que el estudiante ha de identificar, normalmente recibiendo información sobre la corrección de la respuesta enviada por el estudiante. Sin embargo, también podemos encontrarnos con sistemas de entrenamiento productivo audiovisual (Hardison, 2005); Burleson, 2007), donde el hablante compara su pronunciación con la del hablante nativo auditiva y visualmente, en este último caso gracias a los espectrogramas de su producción del habla. Asimismo, también existen procedimientos de entrenamiento de producción visual solamente (May-Lai Tsui, 2005, p. 2), o incluso software que utiliza videojuegos para facilitar la adquisición de contrastes fonéticos (Lim & Holt, 2011). Por otro lado, los objetivos de cada programa de entrenamiento también son diversos: algunos solamente buscan mejorar la discriminación de un contraste fonético particularmente difícil en una población de hablantes de una determinada lengua materna, otros entrenan todas las vocales de la lengua meta agrupándolas en subgrupos según su dificultad de percepción, etc. Dentro del entrenamiento productivo, a veces se busca mejorar la pronunciación de ciertos fonemas, o comprobar si un determinado tipo de entrenamiento mejora la percepción (Bradlow et al., 1997) o la producción (Akahane-Yamada and McDermott, 1998) de la lengua meta sin entrenar este aspecto de forma explícita, sino a través del entrenamiento de uno de estos aspectos sin contar con el otro.

Por otra parte, la variabilidad en el período e intensidad del entrenamiento también es destacable, pudiendo variar desde menos de 2 minutos (de Vos, 2012) minutos hasta más de 22 horas de entrenamiento (Bradlow et al, 1999), todo ello repartido en sesiones realizadas en un día (de Vos, 2012) o espaciadas en períodos de más mes de duración (Burleson, 2007).

2. OBJETIVOS

Realizar una revisión sistemática de los artículos de investigación sobre entrenamiento perceptivo para así esclarecer la calidad de los estudios existentes, y sus principales problemas y fortalezas. Esta revisión incluirá:

- Un análisis estadístico de las principales características de los estudios (a saber: edad media, rango de edad, lengua del sujeto, lengua a aprender, etc.).
- La aplicación de la escala metodológica MQS, de la que se obtiene una puntuación de la calidad de los estudios de acuerdo con criterios metodológicos.
- Identificar los principales problemas metodológicos de estos estudios, basándonos en la escala metodológica y en el análisis estadístico
- Proponer un nuevo modelo de entrenamiento fonético que subsane en la medida de lo posible los principales problemas detectados.

3. METODOLOGÍA

3.1. Instrumentos

Programas informáticos: usamos el software SPSS 20 para los análisis de las principales variables que caracterizan los artículos sobre entrenamiento fonético, así como para las variables de la escala de calidad metodológica utilizadas para codificar los estudios seleccionados.

Escala MQS: constituye un instrumento que evalúa la calidad metodológica de estudios primarios originalmente diseñada para realizar meta-análisis, construida basándose en Sanduete (2008). Se trata de una escala que recoge los ítems citados con más frecuencia para medir la calidad metodológica hasta el año 2010 (Sanduete, 2008), cuya validez de contenido ha sido evaluada por un juicio de expertos (Sanduete, 2008).

3.2. Procedimiento

Búsqueda en bases de datos: llevamos a cabo la búsqueda de información bibliográfica en el área del entrenamiento fonético en el aprendizaje de segundas lenguas en 9 bases de datos informatizadas: SAGE, SCOPUS, Springer LINK, Web of Science, Wiley Online Library, Oxford Journals, Cambridge Journals, Google Scholar, Science Direct. Se obtuvieron un total de 66 estudios, de los cuales se seleccionaron aquellos que trataran los efectos de los programas de entrenamiento perceptivos para el aprendizaje de segundas lenguas, excluyendo aquellos que estuvieran relacionados con el aprendizaje de la lengua materna y los que solamente constituían artículos técnicos presentando nuevos sistemas de entrenamiento de la pronunciación. En total se obtuvieron 26 (el 39%), teniendo en cuenta los criterios anteriores para la realización de este estudio. Los estudios seleccionados fueron publicados desde el año 1984 hasta diciembre de 2012. Las palabras clave utilizadas fueron “pronunciation training”, “perception training”, “phonetic training”, “phonological training” y “production training”. Se recopilieron 26 estudios publicados en revistas, recogidos en tesis doctorales, trabajos de fin de máster y trabajos de fin de grado.

Se codificaron los estudios teniendo en cuenta los 22 ítems de la escala considerados en la escala de calidad metodológica, cuyos ítems se refieren a características extrínsecas (un ítem), sustantivas (6 ítems: 2 relacionados con la muestra, 1 con el contexto y 3 con el tratamiento) y metodológicas (15 ítems). Asimismo, se codificaron las características principales de los estudios para realizar un análisis descriptivo de éstos y poder compararlos con el análisis descriptivo de las variables de calidad metodológica. Estas variables fueron el número de sujetos, la edad media, los criterios de inclusión en cada estudio, el tipo de entrenamiento

fonético, la lengua materna de los sujetos, el nivel de estudios, la existencia de problemas auditivos, la mejora en la producción y/o percepción de la lengua meta (a saber, la lengua extranjera a aprender), la lengua meta, la intensidad y el período de realización del tratamiento.

4. RESULTADOS

4.2. Características de los estudios

De la muestra de 26 estudios analizados, la variable número de sujetos tiene una desviación típica de 20,33 y una media de 25,65 por lo que el grado de dispersión en esta variable es alto, lo que indica la dificultad de comparar los resultados entre estudios, y arroja dudas sobre la robustez de las conclusiones del estudio para la población. La distribución presenta una simetría ligeramente positiva (2,186) con un error típico de asimetría de 0,456, por lo que los valores extremos tienden a estar por encima de la media. La distribución es acusadamente leptocúrtica (6,652), lo que indica que los datos se agrupan en las colas más que en una distribución normal. El 50 por ciento de los estudios tienen un número de sujetos de entre 10,75 y 30,25.

En cuanto a la edad media, la mitad de los datos se han perdido, por lo que quedan solamente trece observaciones. En este caso la desviación típica es de 5,67, por lo que los datos se desvían menos que en la variable anterior. La asimetría es ligeramente negativa (-0,269): esta variable tiene una distribución prácticamente simétrica. La curtosis sigue siendo leptocúrtica (2,938). El 50 por ciento de los estudios tienen sujetos de una edad media de entre 21,5 y 27 años.

Todos los estudios requieren que el participante no sea hablante nativo de la lengua a adquirir. No tomando en cuenta este dato, un 57,7 por ciento de los estudios de entrenamiento perceptivo no requieren de ningún criterio de inclusión adicional, y el 26,9 solo requiere que los sujetos no hayan residido en el país donde esa lengua es hablada por hablantes nativos por un largo período de tiempo. Esto no siempre debería ser un requisito, porque existen datos (Iverson, 2012) que afirman que el entrenamiento fonético también beneficia a sujetos expuestos a la lengua en estos contextos, por lo que no se debe descartar usarlo como requisito. Sólo un 11,9 por ciento de los estudios requería un test de lengua para verificar el nivel de la lengua a adquirir. Un 7,7 requería un test de lengua además del requisito de no residencia, y solamente un estudio tenía en cuenta variables de rendimiento académico. Como se puede observar, los criterios de exclusión son bastante laxos y no tienen en cuenta variables de tipo cognitivo, e incluso del propio nivel de lengua, que puede ser muy diferente del informe que puedan proporcionar los sujetos, aunque en la mayoría de los estudios ni siquiera se pide que los sujetos informen de su nivel de segunda lengua.

Tipo de entrenamiento. El entrenamiento perceptivo auditivo es el tratamiento más usado para facilitar la adquisición de la fonética de la lengua extranjera (usado en 69,2 por ciento de los casos). En segundo lugar, un tratamiento combinado de entrenamiento perceptivo auditivo y productivo visual se usa en el 15,4 por ciento de los casos, y el resto de los casos (15,4 por ciento) utilizan diferentes subtipos de entrenamiento productivo.

Lengua materna de los sujetos. En esta clase de estudios, el mayor porcentaje de sujetos ha sido el grupo de hablantes nativos de japonés (26,9), seguido de estudios realizados con hablantes de diferentes lenguas en un mismo estudio (23,1 por ciento) y hablantes de neerlandés (11,5 por ciento). Como podemos observar, tres grupos de hablantes nativos concentran más de la mitad de las observaciones, por lo que se puede observar una sobrerrepresentación de 2 grupos de hablantes mientras que un 38,5 por ciento de estudios concentran la mayoría de los sujetos que hablan otros nueve idiomas en total.

Tampoco encontramos que el nivel de estudios esté adecuadamente representado en esta muestra, ya que el 42,3 por ciento de los mismos está compuesto por sujetos que tienen estudios universitarios o superiores, y existe un 46,2 por ciento de estudios donde no se aportan datos sobre el nivel de estudios, por lo que no está nada claro el efecto que pudiera tener esta variable en la mejora después del tratamiento. El 11,5 por ciento de las observaciones indica un nivel de estudios primarios (3,8 por ciento) o secundarios (7,7 por ciento).

Las pruebas de audición tampoco parecen ser variables tenidas en cuenta en la mayoría de los estudios, ya que el 69,2 por ciento de estos no indican que se realizaran pruebas de audición. Sólo un 30,8 por ciento de estudios realizaron una audiometría de tonos puros para verificar que los sujetos poseían una audición normal.

Centrando nuestra atención en los resultados del tratamiento, la inmensa mayoría de estudios informan de un aumento de las capacidades de percepción y/o producción de la lengua extranjera (88,5 por ciento). Un 11,5 observa una falta de mejora o ni siquiera aportan información sobre esta variable.

En su mayoría, la lengua meta objeto de aprendizaje es el inglés (73,1 por ciento), como cabía esperar debido al estatus de esta lengua como lengua franca en la mayoría del planeta. El resto de los idiomas objeto de estudio tienen porcentajes muy similares y tienden a aparecer solamente en una observación.

Por otra parte, si analizamos la distribución de la variable tiempo de entrenamiento medido en número de horas, observamos que la media es 5,317, y la desviación típica es 5,688, lo que sugiere que existe un grado de dispersión muy alto. La asimetría de la muestra es positiva, por lo que las observaciones tienden a ser mayores que la media. Sin embargo, el apuntamiento de la distribución es elevado (curtosis = 2,72), por lo que la distribución tiende a acumular un mayor

número de datos en las colas que en una distribución normal. El 50 por ciento de las observaciones presentaba tratamientos de una duración de entre 1,3874 a 5,75 horas.

La variable período de entrenamiento también presenta una desviación típica elevada (10,711), lo que con una media de 14,11 indica el elevado grado de dispersión de los datos. Esta dispersión no implica la existencia de un elevado grado de asimetría (0,271), que se acerca al valor de 0 más que el resto de las variables, aunque tenga una asimetría ligeramente positiva. El coeficiente de curtosis (-1,211) señala que la distribución es platicúrtica, por lo que existe una menor concentración de datos en los valores centrales. El cincuenta por ciento de las observaciones tenía un tratamiento de entre 3,75 a 21 días.

En cuanto a la distribución de sexos en los experimentos estudiados, encontramos que el 46,2 por ciento de las observaciones no se ajustan al criterio de paridad. Este criterio ha sido definido como la existencia de una proporción entre hombres y mujeres de entre el 40 al 60 por ciento, quedando los valores fuera de ese rango los estudios con proporciones más extremas. El 26,9 por ciento siguiente corresponde a estudios donde si existe una cierta paridad en la muestra de sujetos, y el 26,9 por ciento restante se corresponden con estudios que no aportan los datos, por lo que se consideran perdidos.

4.3. Calidad de los estudios

La escala de Calidad Metodológica (MQS), se compone de un conjunto de variables cualitativas que se suman en una variable cuantitativa final que a su vez se divide en tres intervalos cualitativos (a saber: calidad baja, moderada y alta). Dentro de estos intervalos se incluyen variables que solamente suman puntuación si se cumplen sus requisitos (un determinado número de sujetos en la muestra o si se indica la clase de publicación) y variables que además de sumar puntuación para el cómputo final de calidad también son requisitos para alcanzar un grado de calidad determinado (por ejemplo, si el estudio no es de doble ciego la calidad no podrá nunca llegar a ser alta no importa la puntuación alcanzada). Cada variable puede recibir una puntuación de 1 (si se cumple con el requisito), 0 (si no se cumple), y en algunos casos de 0,5 (si se cumplen algunos aspectos solamente). Procedemos con el análisis cualitativo y cuantitativo cuando proceda de estos estudios según las variables de esta escala: tipo de publicación, rango de edad, edad media, ámbito de intervención, período de tratamiento, intensidad del tratamiento, unidades (en grupo o individual), criterios de inclusión, unidades asignadas aleatoriamente, metodología o diseño, tamaño de la muestra, mortalidad, mortalidad entre grupos, exclusiones posteriores a la aleatorización, período de seguimiento, ocasiones de medida para cada variable, medidas en la prueba previa aparecen en la prueba posterior, variables dependientes estandarizadas, técnicas de control, definición del constructo de resultado, métodos

para la imputación de valores perdidos, valor del tamaño del efecto, valoración total y nivel de calidad.

Tipo de publicación. El total de los estudios recogidos indican el tipo de publicación de que se trata. Para esta variable no existe una propuesta de mejora. De la misma manera, el ámbito de intervención (el campo de la educación) siempre se indica, por lo que tampoco se encuentran debilidades en este indicador. Lo mismo ocurre con indican **criterios de inclusión**, ya que todos los estudios tienen al menos un criterio de inclusión. No hubo ninguna **exclusión posterior a la aleatorización** en los estudios que hubo aleatorización luego esta variable tampoco presenta debilidades.

En cuanto al **rango de edad**, el 84,6 por ciento de los estudios seleccionados la indican, por lo que podemos concluir que es una variable altamente controlada en este tipo de estudios. No puede decirse lo mismo de la variable **edad media**, en la que se indica en un porcentaje de 65,4 por ciento, sensiblemente menor que la variable anterior. El 34,6 por ciento de estos estudios no indican la edad media de sus sujetos, dejando la puerta abierta a un mayor sesgo de los datos en caso de que existan sujetos con valores extremos de edad.

Quizás uno de los mayores problemas de estos estudios en lo que a su calidad metodológica se refiere es el **período de tratamiento** de estos estudios, que siempre es inferior o igual a seis meses (en realidad es siempre menor de los cuatro meses), un período insuficiente para obtener datos de mejora de mayor calidad.

Por el contrario, todos los estudios indican el número de sesiones realizadas dentro del entrenamiento, lo que nos indica que no es necesario proponer una mejora en la variable **intensidad del tratamiento**. Lo mismo ocurre con la variable **unidades (en grupo o individual)**, en la que siempre se indica que la intervención se aplicó al grupo de forma individual.

La variable **unidades asignadas aleatoriamente** también tiene ciertas debilidades. Si bien es cierto que el 46,2 por ciento de los estudios tiene algún control de variables extrañas (la edad, la audición, etc...), solamente el 34,6 por ciento tiene una asignación de unidades aleatoria, y hay un porcentaje considerable de casos en los que no hay ningún control de variables extrañas tampoco (19,2 por ciento de los casos).

Hablando de la **metodología o diseño de los estudios**, el 34,6 por ciento de los mismos presenta un diseño preexperimental, ya que no poseen un grupo de control propiamente dicho, o lo tienen, pero no han realizado una medida pre-test. Es escaso el número de estudios con diseño experimental (15,4 por ciento), ya que la mayoría de los estudios carece de una asignación aleatoria de las unidades a las condiciones de intervención (50 por ciento), por lo que se corresponden con casos de diseño cuasi-experimental.

El **tamaño de la muestra** tiene una calidad moderada en la mayoría de los casos (50 por ciento), por tener un número de sujetos de entre 15 y 30. Los casos de calidad baja (menos de 15 sujetos) y alta (más de 30 sujetos) en esta variable se distribuyen de forma bastante equitativa (26,9 y 23,1 por ciento de las observaciones, respectivamente).

La **mortalidad experimental** suele ser baja en la mayoría de los estudios (no hay mortalidad o esta es menor del 30 por ciento en el 84,6 por ciento de las observaciones). Solamente el 15,4 por ciento poseen tasas de mortalidad mayores del 30 por ciento. En estrecha relación con esta variable, la variable **mortalidad entre grupos** presenta un patrón similar al de la variable anterior, ya que la mayoría de los estudios no disponen de grupo de control, y al ser estos los que han presentado mayor mortalidad experimental, la mortalidad entre grupos es homogénea en un 92,3 por ciento de los casos.

Teniendo en cuenta la variable **período de seguimiento**, está podría mejorarse en todos los casos, ya que en ningún caso se hizo un seguimiento superior a seis meses, e incluso los seguimientos mayores de un mes son raros en la muestra. Por tanto, el mantenimiento de las mejoras producidas por los programas de entrenamiento no se ha demostrado a tan largo plazo en esta muestra.

En las **ocasiones de medida para cada variable** tampoco se observan debilidades: en todos los estudios analizados se obtienen medidas antes y después del periodo de intervención, a excepción de variables como el nivel de gramática del inglés, que se piensa que es muy estable para períodos tan reducidos de tiempo. Esta variable presenta un comportamiento muy similar a la **de medidas en la prueba previa aparecen en la prueba posterior**, solamente hubo un 3,8 por ciento de estudios (un solo estudio), donde las medidas de la prueba previa no aparezcan en la prueba posterior.

Criterio de **Variables dependientes estandarizadas**. Aunque muy pocos estudios utilizaron herramientas estructuradas, la gran mayoría utilizó un procedimiento homogéneo, y únicamente el 7,7 por ciento de los datos pertenecen a estudios realizados con herramientas no estructuradas. Por otra parte, la variable **técnica de control** muestra gran debilidad: el 96,2 por ciento de los estudios no aplica ninguna técnica de control, y el único que aplica alguna técnica simplemente se trata de un estudio ciego. No hay ningún estudio de doble ciego en la muestra.

Por el contrario, la mayoría de los estudios proveen una **definición conceptual o empírica del constructo** de las variables dependientes de forma que es replicable por el lector (88,5 por ciento), y solamente un 7,7 por ciento define al menos una variable de forma conceptual o empírica, por lo que esta variable no presenta una gran debilidad según la escala aplicada, a excepción de un estudio (3,8 por ciento) que no lo define.

Imputación de valores perdidos. No encontramos valores perdidos en ninguno de los datos, por lo que la variable imputación de los valores perdidos siempre suma puntos a los estudios, al no quedar valores perdidos sin imputar. Esto parecería una contradicción al existir mortalidad en algunos estudios, pero en esos casos la mortalidad se produce en la fase de selección, porque en estas observaciones no los sujetos que abandonan el tratamiento ni siquiera lo comienzan. Otra duda que podría surgir es porque en algunas variables de las características mencionadas anteriormente presentan valores perdidos. Esto ocurre porque muchos estudios no tienen en cuenta variables de potencial influencia que, si se tienen en cuenta en esta revisión, como la proporción de hombres y mujeres. Asimismo, todos los estudios especifican el **valor del tamaño del efecto**.

Sumando la **valoración total** de todas las variables analizadas anteriormente en la escala, encontramos que casi todos los estudios alcanzan una calidad moderada. La media es de 16,115, y la desviación típica es de 1,710, lo que indica que los datos se concentran alrededor de la media. La distribución presenta una asimetría negativa (-1,252), por lo que la muestra tiende a acumular más datos inferiores a la media. El coeficiente de Curtosis tiene un valor de 3,535. Es decir, la distribución presenta un apuntamiento muy elevado.

4.4. Fortalezas y posibles mejoras

a) Fortalezas

La elevada desviación típica en la variable **número de sujetos** es uno de los problemas más importantes. Alrededor de un cuarto de los estudios analizados presenta una calidad baja en esta variable (menos de 15 sujetos). Esto afecta en gran medida a la representatividad de los datos, y dado que el tratamiento no presenta riesgos de ningún tipo para los sujetos, incrementar el tamaño de la muestra hasta un número de entre 25 y 30 no presentaría problemas. En cuanto a la variable edad media, el principal problema es que no se menciona en el 50 por ciento de los estudios. La recomendación en este caso es siempre aportar datos de la edad media.

Los **criterios de inclusión y exclusión** destacan por ser escasos y a menudo no incorporar pruebas que controlen variables extrañas. Dado que la mayoría de estudios solamente requieren que el hablante no sea hablante nativo de la lengua a entrenar, otras variables como el nivel de segunda lengua a entrenar podrían sesgar los resultados. Para mitigar estos problemas, deberían de poder realizarse pruebas de aptitud para aprender segundas lenguas como la prueba de aptitud para lenguas extranjeras de Pimsleur, o pruebas de memoria de trabajo o atención sostenida, que han demostrado tener influencia en el aprendizaje de la lengua (Miyake & Friedman, 1998; Finneran et al, 2009).

En la variable **tipo de entrenamiento**, existen pocos estudios realizados con entrenamiento productivo, probablemente por la dificultad que requiere enseñar a los sujetos cómo interpretar las herramientas de entrenamiento productivo más común, que es el espectrograma y la onda sonora. Una combinación de entrenamiento productivo y perceptivo también se observa en una minoría de los resultados, probablemente por las mismas razones. Dado que todavía no está claro qué ventajas produce cada tipo de entrenamiento, sería una buena recomendación realizar un programa de entrenamiento con grupos de entrenamiento productivo, perceptivo y combinado por separado, para evaluar los efectos en la producción y la percepción de la lengua extranjera por separado, con un grupo de control que no tuviera tratamiento, para esclarecer qué tratamiento es más efectivo y qué parámetros mejora cada tipo de entrenamiento.

Asimismo, la variable **lengua materna** de los sujetos es de difícil análisis, pues es una decisión que depende del tipo de datos que necesitemos obtener. En el caso de que queramos obtener datos sobre el efecto del tipo de entrenamiento de forma específica en un determinado grupo de hablantes, entonces es importante tener un grupo procedente de un solo grupo de lengua. Por otro lado, si queremos obtener información sobre los efectos generales del efecto perceptivo en hablantes de otras lenguas, deberíamos incluir un grupo de hablantes de segundas lenguas variado. Como este supuesto es poco práctico (se necesitarían cientos de miles de sujetos para tener un estudio que incluyera a los hablantes de las principales lenguas del mundo), la propuesta es realizar un estudio con hablantes de la misma lengua.

Teniendo en cuenta el **nivel de estudios**, nos encontramos con que el 46,2 por ciento no aporta datos en este sentido, por lo que prácticamente no se tiene en cuenta como factor que influya en la percepción. De los datos existentes, existe una sobrerrepresentación de los sujetos con estudios universitarios (42,3 por ciento), probablemente porque los investigadores se encuentran afiliados a instituciones universitarias en la mayoría de los casos. La propuesta para esta variable sería un control de esta variable en los cuestionarios de inclusión, que apenas se controla. Por otra parte, existe una cierta escasez de estudios de entrenamiento en grupos que hayan cursado estudios secundarios o primarios, quizás por la mayor dificultad de los investigadores a seleccionar grupos fuera de su ámbito de trabajo. Por ello, y a pesar de que esto no afecta a la estructura del programa de entrenamiento en sí, más investigaciones en este grupo de edad podría aportar más información sobre los efectos de este tipo de entrenamiento en la población.

De la misma forma que en la variable lengua materna, no hay propuestas mejoras específicas para la variable **lengua meta**. Sin embargo, debido a la hegemonía del inglés en el estudio de la adquisición de lenguas, no hay muchos estudios que se ocupen de analizar el efecto del entrenamiento fonético sin tener en cuenta la lengua objeto, y ya que se necesitarían miles de sujetos para realizar un estudio con

diferentes lenguas meta, es más sensato mantener una sola lengua como lengua meta, ya sea el inglés u otra lengua.

Este no es el caso en la variable **tiempo de entrenamiento**, que tiene una dispersión muy elevada, con estudios que varían desde menos de 2 minutos de entrenamiento hasta más de 20 horas. Por ello, no se conoce con seguridad el tiempo necesario para producir unas mejoras determinadas en la percepción y/o producción de fonemas de la lengua extranjera. En este caso, y encontrando que algunos estudios donde no hay mejoras se caracterizan por tener períodos de entrenamiento muy cortos, recomendamos que el número de horas se aproxime al máximo posible, en torno a las 20 horas de entrenamiento para períodos de entrenamiento cortos (un mes), hasta 60 horas para períodos de entrenamiento más largos (6 meses), de manera que se disponga de al menos de 2 horas de entrenamiento semanal.

Comentarios similares se le pueden hacer a la variable **período de tratamiento**, cuya desviación típica es también muy elevada. En este caso, sería recomendable que el tratamiento se extienda hasta períodos de al menos un mes, si queremos evitar resultados negativos como los mencionados en la variable tiempo de entrenamiento.

En cuanto a la existencia de pruebas de **audición** en los programas de entrenamiento auditivo, esta destaca por su ausencia en el 69,2 por ciento de los casos. Esto presenta un gran problema a la hora de evaluar los efectos de estos programas, ya que una percepción disminuida puede producir resultados peores en el entrenamiento, y no se controla la posible existencia de personas que tienen problemas de audición pero no lo saben: en este caso el auto-informe no es fiable.

La falta de control en un gran número de casos, así como la alta dispersión en el número de sujetos nos hace dudar de los efectos del entrenamiento, que demuestra ser positivo en el 88,5 por ciento de los casos.

La variable genera constituye un punto a mejorar dentro de los estudios seleccionados. No solamente es escaso el número de estudios que entran dentro del criterio de paridad (incluso definiendo la paridad de una manera relativamente laxa, con una proporción de entre el 40-60 por ciento), sino que también existe un elevado porcentaje de estudios que no aportan datos sobre el género de los sujetos, por lo que ni siquiera consideran que esta variable puede ejercer una influencia en la variable mejora de la percepción y/o producción. La mejora que se plantea en este caso es simplemente la recogida de una mayor muestra que permita que haya paridad entre hombres y mujeres.

b) Resultados de la aplicación de la escala MQS

En este apartado solamente comentaremos aquellas variables que no hayan obtenido la puntuación máxima en todas las observaciones analizadas. Por tanto, de los siguientes ítems no propondremos mejores en la calidad de sus variables, ya que

han demostrado tener la máxima calidad en todos los estudios seleccionados: tipo de publicación, intensidad de tratamiento, unidades (en grupo o individual), criterios de inclusión y exclusión de las unidades proporcionados, exclusiones posteriores a la aleatorización, ocasiones de medida para cada variable, imputación de valores perdidos y valor del tamaño del efecto.

A pesar de que se menciona en la mayoría de las observaciones, el **rango de edad** no aparece en todos los estudios, y constituye una debilidad menor, pero a tener en cuenta en un programa de entrenamiento. Un mayor problema en la calidad metodológica es el período de tratamiento, que requiere de **un período de tratamiento** mayor de 6 meses para asignar la mayor calidad metodológica al estudio. Este período es menor en todos los estudios observados. Asimismo, la variable **unidades asignadas aleatoriamente** a condiciones obtiene una calidad moderada por no controlar solamente algunas variables extrañas sin hacer uso de la asignación aleatoria excepto en el 34,6 por ciento de las observaciones. La recomendación en este caso es la asignación aleatoria de unidades siempre y cuando se dispongan de suficientes sujetos en cada grupo (para una muestra total de 30 sujetos o más).

En cuanto al **diseño o metodología** de los estudios, nos encontramos con que solamente un 15,4 por ciento de ellos son experimentales, siendo uno de los principales problemas a la hora de asignar la máxima calidad a esta variable. La mayoría no utilizan la asignación aleatoria, aunque la calidad de la mayoría de los estudios es moderada para esta variable ya que usan métodos de recogida de datos semiestructurados. La propuesta de mejora que haríamos sería utilizar grupos de control y asignación aleatoria de los sujetos a grupos.

Hablando del **tamaño de la muestra**, lo ideal sería encontrar un número de sujetos mayor de 30, aunque no es una variable tan importante si la comparamos con la falta de control de otras variables mencionadas anteriormente, así como la falta de aleatorización. Las variables **mortalidad experimental** y **mortalidad entre grupos** tienen una calidad alta por ser la mortalidad baja en ambos, por lo que no hay recomendaciones especiales en este caso. No mencionaremos tampoco la variable **medidas en la prueba previa aparecen en la prueba posterior**, ya que solamente un estudio no cumple con los criterios de esta variable. Algo similar ocurre con la variable **definición de constructo de resultado**, así que tampoco se ofrecen propuestas de mejora en este caso.

Por el contrario, la variable **período de seguimiento** es siempre menor a lo establecido como criterio de calidad en la escala, por lo que sería necesaria aumentar el seguimiento a más de 6 meses, aunque lo ideal sería hacerlo de más de 12 meses para obtener la máxima puntuación en esta área.

Otra gran debilidad identificada en la escala de calidad son las **técnicas de control**. De hecho, es la variable que limita la calidad de la gran mayoría de estudios que obtendrían la calificación de calidad alta si no fuera porque no incorporan

ninguna técnica de control. Dada la dificultad de realizar un estudio de doble ciego, la incorporación de una técnica de ciego sería la propuesta para esta variable, pues no incorpora tantas dificultades y sería suficiente para alcanzar el siguiente nivel de calidad. Esta variable constituye uno de los requisitos limitantes para alcanzar la calidad alta de los estudios, por lo que sería una prioridad por encima de la mejora de otras variables.

5.MODELO DE ENTRENAMIENTO METODOLÓGICO

A continuación, proponemos un modelo de entrenamiento que tenga en cuenta las debilidades encontradas en estos estudios, tratando de incluir en la medida de lo posible las propuestas anteriormente comentadas.

5.1. Número de sujetos y control de variables extrañas

El programa de entrenamiento contará con un grupo mayor de 30 sujetos (40 sería un número adecuado). De este número, la mitad serían asignados a un grupo de control al que solo se realizarían las pruebas iniciales de pre-test y post-test sobre percepción y producción fonética y la recogida de datos para controlar variables de influencia (edad, audición, trastornos cognitivos, nivel de estudios, aptitud lingüística, nivel de lengua anterior, puntuación en atención sostenida y memoria de trabajo). De estas variables potencialmente confundidas, será obligatorio realizar las pruebas correspondientes para todas las variables excepto para las variables cognitivas de atención sostenida y memoria de trabajo, ya que al estar relacionadas con la aptitud lingüística sus efectos son en parte predichos por el test de aptitud en lo que a la adquisición de lenguas se refiere.

5.2.El género requiere de comentarios más específicos como variable de influencia

Este modelo requiere de un reclutamiento de sujetos que trate de controlar la paridad de sexos, excluyendo parte de los datos si la muestra recogida es mayor de 30 a cambio de que exista paridad. Si la muestra es menor de treinta sujetos habría que ser más cuidadoso, pues podríamos perder calidad en la variable de calidad tamaño de la muestra, además de la propia representatividad de los datos. En casos extremos donde la muestra sea baja, mantendremos toda la muestra y aceptaremos el sesgo posible del género a cambio de no tener una muestra aún menor.

5.3. Edad media y Rango de Edad

No solamente son requisitos que se deben incluir como variables que podrían afectar a los efectos del tratamiento, sino que también habría que incluir información

relativa al grado de dispersión de los datos en la muestra. Por otra parte, podría evitarse la inclusión de medidas de dispersión si se observa que el rango de edad es reducido, algo que es discutible en este caso, ya que los estudios tienen rangos de edad desde los 11 hasta más de 40 años, por lo que sería posible que existiera una desviación típica muy elevada y la interpretación de los datos finales sería más complicada. Asimismo, la edad debería tener un rango reducido (preferiblemente de 5 años) ya que es un factor que puede influir en el efecto del entrenamiento.

5.4. Criterios de inclusión y exclusión de los sujetos

Entre los criterios de inclusión, los requisitos deberían incluir ser monolingüe dependiendo de la lengua materna de la población objetivo que adquiere la segunda lengua, así como limitar el estudio a la adquisición de una segunda lengua sin utilizar varios idiomas en el mismo estudio. Asimismo, el requisito de no residencia debería mantenerse por ahora, o al menos controlarse, ya que el elevado número de horas de exposición a la lengua meta en un país donde la lengua se habla por nativos puede influir en el resultado

5.5. Tipo de entrenamiento

El entrenamiento perceptivo es el tratamiento más usado por su sencillez de implementación (no se necesita instruir a los sujetos) y creación (solamente síntesis, modificación e introducción de estímulos en un software de presentación de estímulos). Todavía no está claro si este tipo de entrenamiento es más efectivo que el entrenamiento productivo, o si la combinación de ambos es más efectiva que las anteriores. Sin embargo, dado que la mayoría de estudios se han realizado en entrenamiento perceptivo, hay más datos sobre su efectividad, y por el momento, sería el tipo de entrenamiento recomendado en esta propuesta.

5.6. Lengua materna

La lengua materna siempre será uniforme en cada estudio.

5.7. Tiempo y período de entrenamiento.

Para alcanzar los criterios de calidad del período de entrenamiento, debería situarse en torno a 2 horas por semana para un tratamiento de algo más de seis meses, con un total de unas 48 horas a lo largo de este tiempo. A pesar de que este nuevo modelo duplica el número de horas del estudio más amplio de la muestra (22.5 horas), este último estudio se realizó en cuatro semanas, por lo que la intensidad del tratamiento fue alta, aunque no se alcanzó este criterio de calidad. Debido a

la alta mortalidad experimental en tratamientos tan largos, no se recomienda necesariamente que el tratamiento se extienda más de tres meses a menos que se consiga una mortalidad baja, ya que otras medidas de la escala pueden hacer subir la puntuación y darle alta calidad al estudio.

5.8. Unidades asignadas aleatoriamente

Además de un control de variables extrañas, la asignación de unidades aleatoriamente es fundamental, por el requisito limitante de calidad que le da a esta variable. Por ello, los grupos deben ser lo suficientemente equilibrados para controlar las variables independientes antes mencionadas, especialmente en el caso de que no se puedan asignar las unidades aleatoriamente.

5.9. Metodología o diseño

En estos grupos de población, debe de asignarse un grupo de control en todos los casos (diseño experimental), o al menos un grupo control no equivalente, en un diseño cuasi experimental (por ejemplo, en este caso hablantes nativos de la lengua meta).

5.10. Período de seguimiento

Dado que un seguimiento de más de un año podría suponer costes elevados, un seguimiento de algo más de seis meses tendría la mitad de la puntuación y supondría costes menores a la hora de realizar el programa, y esta es la solución que proponemos.

5.11. Variables dependientes estandarizadas

No existen todavía pruebas de percepción fonética para el inglés que se hayan validado y estén ampliamente extendidas en la enseñanza del inglés como segunda lengua. Asimismo, la evaluación de la pronunciación de la lengua no se ha podido estandarizar en un test, y todavía se realiza a través de la puntuación de jueces nativos, ya sea fonetistas expertos o no. Por ello nuestra propuesta consiste en usar procedimientos homogéneos en las medidas pre- y post- test, como usan la mayoría de los estudios mencionados.

5.12. Técnicas de control

Esta variable es el principal problema encontrado en los estudios, y la principal barrera para que la mayoría de ellos alcancen la nota más alta en la escala. La solución más rápida parece ser implementar un enmascaramiento en forma de un

estudio ciego, en la que los sujetos no sepan el fin exacto del tratamiento, aunque esto en realidad es difícil de conseguir por completo, ya que el tratamiento requiere realizar tareas lingüísticas que, aunque no se proporcione información sobre los objetivos de estas, sería fácil para los sujetos deducir que tiene el objetivo de mejorar la adquisición de la lengua extranjera.

5.13. Puntuación total

Si se siguieran estas directrices, además de las que ya son comunes en la mayoría de los estudios de entrenamiento fonético, la puntuación total sería de 19, y la calidad de este programa de entrenamiento sería alta de acuerdo con la escala de calidad metodológica MQS.

6. CONCLUSIONES Y DISCUSIÓN

Tras el análisis estadístico, se demuestra que los estudios de entrenamiento fonético para el aprendizaje de segundas lenguas pueden mejorarse en diversos aspectos, a pesar de cumplir con la mayoría de los ítems en la escala de calidad metodológica usada. Las principales fortalezas de estos estudios en cuanto a sus calidad metodológica consisten en la escasa mortalidad y mortalidad entre grupos, el uso de procedimientos homogéneos para la recogida de datos antes y después de las intervenciones, la indicación del tamaño del efecto y la existencia de pruebas antes y después del tratamiento, además de la adecuada definición del constructo teórico, la indicación de criterios de inclusión y exclusión, y la indicación de la intensidad del tratamiento. Por otro lado, las mayores debilidades se encuentran en variables cuantitativas que escapan a la medida de la escala de calidad usada, como el tamaño de la muestra, que es muy variable y a veces demasiado pequeño para ser representativo de la población, el rango de edad de los sujetos, que no se menciona en muchos casos, así como el uso de períodos y duración de tratamiento muy dispares y a menudo reducidos. Asimismo, existen variables cuyos parámetros no se especifican generalmente a pesar de su potencial influencia en los resultados, como el nivel de estudios o el nivel de audición. En lo que respecta a la escala, la falta de período de seguimiento adecuado, un tamaño de la muestra suficientemente grande, y, sobre todo, la falta de cualquier técnica de control en la inmensa mayoría de los estudios constituye los principales impedimentos para que los estudios tengan la máxima calidad.

Basándonos en este análisis, las propuestas de mejora se han enfocado a ampliar el período, intensidad y seguimiento del tratamiento, el uso de la forma más sencilla de enmascaramiento y el control de variables que influyen en los resultados. En conclusión, pretendemos que este estudio constituya una futura recomendación para obtener resultados más fiables en los estudios de entrenamiento fonético que se realicen en estudios posteriores en el área de la Psicolingüística.

7. BIBLIOGRAFÍA

- Akahane-Yamada, R., McDermott, E., Adachi, T., Kawahara, H., & Pruitt, J. S. (1998). Computer-based second language production training by using spectrographic representation and HMM-based speech recognition scores. *Proceedings of ICSLP*, Sydney, Australia.
- Bradlow, A. R., Pisoni, D.B., Yamada, R.A., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America*, 101, 2299–2310.
- Bradlow, A.R., Akahane-Yamada, R., Pisoni, D.B., & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception & Psychophysics*, 61, 977–985. <http://faculty.wcas.northwestern.edu/ann-bradlow/publications/1999/bradlow-et-al-r15-p&p99.pdf>.
- Bergström, H. (2007). *Evaluation of a computer assisted language learning system for Swedish language learners*. (Unpublished M. Sc. Thesis). Royal Institute of Technology, Stockholm.
- Burleson, D.F. (2007). Improving intelligibility of non-native speech with computer assisted phonological training. *IULC Working Papers Online*. <https://www.indiana.edu/~iulcwp/pdfs/07-Burleson5.pdf>.
- Cucchiari, C., Neri, A., & Strik, H. (2009). Oral proficiency training in Dutch L2: The contribution of ASR-based corrective feedback. *Speech Communication*, 51, 853–863.
- De Vos, J. (2012). *Does enhanced bimodal distributional training improve perception of English /æ/ and /ɛ/ for adult native speakers of Dutch?* (Unpublished B.A. thesis) University of Amsterdam, Amsterdam.
- De Wilde, E. (2010). Perception and production in second language phonology: The effect of audiovisual training on the acquisition of the English dental fricatives. (Unpublished B.A. thesis) University of Gent, Gent.
- Finneran, D., Francis, A., & Leonard, L. (2009). Sustained attention in children with specific language impairment. *Journal of Speech, Language and Hearing Research*, 52, 915–929.
- Hardison, D.M. (2005). Second-language spoken word identification: Effects of perceptual training, visual cues, and phonetic environment. *Applied Psycholinguistics*, 26, pp. 579–596. doi: 10.1017.S0142716405050319.
- Heeren, W.F.L., & Schouten, M. E. H. (2010). Perceptual development of the Finnish /t-t:/ distinction in Dutch 12-year-old children: A training study. *Journal of Phonetics*, *Journal of Phonetics*, 38(4), 594–603. <http://dx.doi.org/10.1016/j.bbr.2011.03.031>.
- Herd, W. (2011). *The perceptual and production training of in L2 Spanish: Behavioural, psycholinguistic, and neurolinguistic evidence*. (Unpublished doctoral dissertation). University of Kansas, Lawrence.

- Iverson, P. & Evans, B.G. (2007). Auditory training of English vowels for first language speakers of Spanish and German. *In: 16th International Congress of Phonetic Sciences*, pp. 1625-1628, Saarbrücken, Germany.
- Iverson, P., Pinet, M., & Evans, B. G. (2012). Auditory training for experienced and inexperienced second-language learners: Native French speakers learning English vowels. *Applied Psycholinguistics*, 33 (2012), 145–160. doi:10.1017/S0142716411000300.
- Kaan E, Barkley C., Wayland R., Bao M. (2007). Effects of native Language and training on lexical tone perception: an ERP study. *Brain Res 1148*, 113-122.
- Kaan E., Barkley C. M., Bao M., Wayland R. (2008). Thai lexical tone perception in native speakers of Thai, English and Mandarin Chinese: An event-related potentials training study. *BMC Neurosci.* 9, 1–17. doi: 10.1186/1471-2202-9-53.
- Kondaurova, M., and Francis, A. (2010). The role of selective attention in the acquisition of English tense and lax vowels by native Spanish listeners: Comparison of three training methods. *J. Phonetics* 38, 569–587.
- Kröger B.J., Birkholz, P., Hoffmann, R., & Meng, H. (2010). Audiovisual tools for phonetic and articulatory visualization in computer-aided pronunciation training. In A. Esposito et al. (Eds.), *Development of Multimodal Interfaces: Active Listening and Synchrony* (pp. 337-345). Berlin: Springer-Verlag.
- Lengeris, A. (2008). The effectiveness of auditory phonetic training on Greek native speakers' perception and production of Southern British English vowels. *Proceedings of ISCA Tutorial and Research Workshop On Experimental Linguistics*. Athens, Greece: International Speech Communication Association.
- Lim, S. J., & Holt, L. L. (2011). Learning foreign sounds in an alien world: Videogame training improves non-native speech categorization. *Cognitive Science*; 35(7), 1390–1405.
- Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America*, 94, 1242–1255.
- Miyake, A. & Friedman, N.P. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In Healy, A. & Bourne, L. (eds.), *Foreign Language Learning: Psycholinguistics Studies on Training and Retention*. Mahwah, NJ: Lawrence Erlbaum.
- Nishi, K., & Kewley-Port, D. (2007). Training Japanese listeners to perceive American English vowels: Influence of training sets. *Journal of Speech, Language, and Hearing Research*, 50, 1496–1509.
- 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. <http://dx.doi.org/10.1080/09588220802447651>.

- Nishi, K. & Kewley-Port, D. (2007). Training Japanese listeners to perceive American English vowels: Influence of training sets. *Journal of Speech, Language and Hearing Research*, 50 (6), 1496–1509.
- Nishi, K., & Kewley-Port, D. (2008). Nonnative speech perception training using vowel subsets: Effects of contrast choice and order of training. *Journal of Speech, Language and Hearing Research*, 51 (6), 1480-1493.
- Sanduvete Chaves, S. (2008). Innovaciones metodológicas en la evaluación de la formación continua. (Unpublished doctoral dissertation). Universidad de Sevilla, Sevilla.
- Strange W, Dittmann S. (1984). Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Percept Psychophys.*, 36, 131–145.
- Tsui, H.M.L. (2012). Ultrasound speech training for Japanese adults learning English as a second. (Unpublished M.Sc. dissertation). The University of British Columbia, Vancouver.
- Thomson, R.I. (2012). Improving L2 listeners' perception of English vowels: a computer-mediated approach. *Language Learning*, 62:4, 1231–1258. <http://dx.doi.org/10.1111/j.1467-9922.2012.00724.x>
- Ylinen, S., Uther, M., Latvala, A., Vepsäläinen, S., Iverson, P., Akahane-Yamada, R., & Näätänen, R. (2009). Training the brain to weight speech cues differently: a study of Finnish second-language users of English. *Journal of Cognitive Neuroscience* 22(6), 1319–1332.

ANEXOS

1. Tabla de resultados de las características de los estudios

1.1 Variables cualitativas

Tipo de tratamiento		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	perceptivo auditivo	18	69,2	69,2	69,2
	perceptivo audiovisual	1	3,8	3,8	73,1
	productivo visual	1	3,8	3,8	76,9
	combinado	4	15,4	15,4	92,3
	productivo audiovisual	1	3,8	3,8	96,2
	productivo auditivo	1	3,8	3,8	100,0
	Total	26	100,0	100,0	

Lengua materna		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	inglés	1	3,8	3,8	3,8
	japonés	7	26,9	26,9	30,8
	checo	1	3,8	3,8	34,6
	español	1	3,8	3,8	38,5
	chino	2	7,7	7,7	46,2
	neerlandés	3	11,5	11,5	57,7
	griego	1	3,8	3,8	61,5
	italiano	1	3,8	3,8	65,4
	francés	1	3,8	3,8	69,2
	coreano	1	3,8	3,8	73,1
	varios	6	23,1	23,1	96,2
	finés	1	3,8	3,8	100,0
	Total	26	100,0	100,0	

Nivel de estudios		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	NAD	12	46,2	46,2	46,2
	primarios	1	3,8	3,8	50,0
	secundarios	2	7,7	7,7	57,7
	universitarios	11	42,3	42,3	100,0
	Total	26	100,0	100,0	

Realización de prueba de audición		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	NAD	17	65,4	65,4	65,4
	no	9	34,6	34,6	100,0
	Total	26	100,0	100,0	

Resultados de la intervención		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	no mejora	1	3,8	3,8	3,8
	mejora	23	88,5	88,5	92,3
	NAD	2	7,7	7,7	100,0
	Total	26	100,0	100,0	

Lengua meta		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	inglés	19	73,1	73,1	73,1
	alemán	1	3,8	3,8	76,9
	holandés	1	3,8	3,8	80,8
	sueco	1	3,8	3,8	84,6
	finés	1	3,8	3,8	88,5
	tailandés	2	7,7	7,7	96,2
	español	1	3,8	3,8	100,0
	Total	26	100,0	100,0	

1.2. Variables cuantitativas

Estadísticos		muestra	Edad media	Tiempo total entrenamiento (horas)	Periodo de entrenamiento (días)
N	Válidos	26	13	24	26
	Perdidos	0	13	2	0
Media		25,65	24,00	5,3179	14,1154
Mediana		24,50	25,00	3,8750	14,0000
Moda		6ª	27	5,00	21,00
Desv. típ.		20,331	5,627	5,68851	10,71196
Varianza		413,355	31,667	32,359	114,746
Asimetría		2,186	-,269	1,663	,271
Error típ. de asimetría		,456	,616	,472	,456
Curtosis		6,652	2,938	2,474	-1,211
Error típ. de curtosis		,887	1,191	,918	,887
Rango		98	25	22,50	35,00
Suma		667	312	127,63	367,00

	10	6,00	14,60	,2150	1,0000
	20	9,40	20,80	1,0000	2,4000
	25	10,75	21,50	1,3875	3,7500
	30	12,30	22,00	1,9000	4,1000
	40	21,00	22,60	2,6000	9,0000
Percentiles	50	24,50	25,00	3,8750	14,0000
	60	27,20	25,40	4,2000	21,0000
	70	29,90	26,80	5,0000	21,0000
	75	30,25	27,00	5,7500	21,0000
	80	33,40	27,00	11,0000	25,2000
	90	48,60	32,40	14,2500	28,6000

a. Existen varias modas. Se mostrará el menor de los valores.

Paridad		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	no existe paridad de sexos	12	46,2	63,2	63,2
	Existe paridad de sexos	7	26,9	36,8	100,0
	Total	19	73,1	100,0	
Perdidos	Sistema	7	26,9		
Total		26	100,0		

2. Tablas de los resultados de la escala de calidad metodológica

2.1 Variables cuantitativas

Tipo de publicación		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	1,00	26	100,0	100,0	100,0

Edad rango		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	,00	4	15,4	15,4	15,4
	1,00	22	84,6	84,6	100,0
	Total	26	100,0	100,0	

Edad media		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	,00	9	34,6	34,6	34,6
	1,00	17	65,4	65,4	100,0
	Total	26	100,0	100,0	

Ámbito de intervención	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
,00	1	3,8	3,8	3,8
Válidos 1,00	25	96,2	96,2	100,0
Total	26	100,0	100,0	

Periodo de tratamiento	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos ,00	26	100,0	100,0	100,0

Periodo de tratamiento	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos ,00	26	100,0	100,0	100,0

Intensidad del tratamiento	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos 1,00	26	100,0	100,0	100,0

Unidades (aplicación de la intervención en grupo o individualmente)	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos 1,00	26	100,0	100,0	100,0

Criterios de inclusión o exclusión	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos 1,00	26	100,0	100,0	100,0

Unidades asignadas aleatoriamente	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
,00	5	19,2	19,2	19,2
Válidos ,50	12	46,2	46,2	65,4
1,00	9	34,6	34,6	100,0
Total	26	100,0	100,0	

Metodología/diseño	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
,00	9	34,6	34,6	34,6
Válidos ,50	13	50,0	50,0	84,6
1,00	4	15,4	15,4	100,0
Total	26	100,0	100,0	

Tamaño de la muestra	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
,00	7	26,9	26,9	26,9
,50	13	50,0	50,0	76,9
1,00	6	23,1	23,1	100,0
Total	26	100,0	100,0	

Mortalidad	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
,00	4	15,4	15,4	15,4
1,00	22	84,6	84,6	100,0
Total	26	100,0	100,0	

Mortalidad entre grupos	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
,00	2	7,7	7,7	7,7
1,00	24	92,3	92,3	100,0
Total	26	100,0	100,0	

Periodo de seguimiento	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
,00	26	100,0	100,0	100,0

Ocasiones de medida en cada variable	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
1,00	26	100,0	100,0	100,0

Medidas de la prueba previa aparecen en la prueba posterior	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
,50	1	3,8	3,8	3,8
1,00	25	96,2	96,2	100,0
Total	26	100,0	100,0	

Variables dependientes estandarizadas	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
,00	2	7,7	7,7	7,7
1,00	24	92,3	92,3	100,0
Total	26	100,0	100,0	

Técnicas de control	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
,00	25	96,2	96,2	96,2
Válidos ,50	1	3,8	3,8	100,0
Total	26	100,0	100,0	

Definición del constructo de resultado	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
,00	1	3,8	3,8	3,8
Válidos ,50	2	7,7	7,7	11,5
1,00	23	88,5	88,5	100,0
Total	26	100,0	100,0	

Imputación de valores perdidos	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos 1,00	26	100,0	100,0	100,0

Valor del tamaño del efecto	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos 1,00	26	100,0	100,0	100,0

Calidad	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
baja	1	3,8	3,8	3,8
Válidos moderada	24	92,3	92,3	96,2
alta	1	3,8	3,8	100,0
Total	26	100,0	100,0	

2.2. Variable cuantitativa

Estadísticos	
Total puntuación de la escala MQS	
N	Válidos 26 Perdidos 0
Media	16,1154
Mediana	16,5000
Moda	16,50
Desv. típ.	1,71060
Varianza	2,926
Asimetría	-1,252
Error típ. de asimetría	,456

Curtosis		3,535
Error típ. de curtosis		,887
Suma		419,00
	10	14,3500
	20	15,0000
	25	15,0000
	30	15,0500
	40	16,4000
Percentiles	50	16,5000
	60	16,5000
	70	16,9500
	75	17,0000
	80	17,3000
	90	18,1500

Paridad entre sexos		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	no existe paridad de sexos	12	46,2	63,2	63,2
	Existe paridad de sexos	7	26,9	36,8	100,0
	Total	19	73,1	100,0	
Perdidos	Sistema	7	26,9		
Total		26	100,0		

