

Similitud fonética en la comparación de los fonemas fricativos del chino y del español

Zhao Liu
East China Normal University (China)  

María Jesús Machuca
Universitat Autònoma de Barcelona (España)  

<https://dx.doi.org/10.5209/clac.87746>

Enviado: 24 de marzo de 2023 • Aceptado: 21 de enero de 2024

ES Resumen: La comparación de los fonemas similares entre lenguas diferentes suele centrarse en un contraste fonémico y son pocos los estudios que han partido de los parámetros acústicos para analizar una similitud fonética entre los fonemas compartidos por dos lenguas. Este estudio intenta comparar los fonemas comunes entre dos lenguas muy diferentes, el chino y el español, a partir del análisis de los parámetros acústicos para determinar si dichos fonemas también muestran unas características acústicas similares. Para llevar a cabo este estudio, dos grupos de informantes, nativos de cada una de las lenguas analizadas, han leído textos que contienen fonemas fricativos en chino mandarín y en español. Se han llevado a cabo comparaciones entre grupos de hablantes con lenguas nativas diferentes y entre los mismos hablantes leyendo en su lengua materna y en una lengua extranjera. Los resultados señalan que, aunque los fonemas son similares en las dos lenguas, los valores de los parámetros acústicos pueden mostrar diferencias. Además, los hablantes chinos han establecido una misma categoría fonética para los fonemas fricativos de las dos lenguas analizadas, la nativa y la extranjera, mientras que los hablantes españoles aún no han igualado esas categorías fonéticas.

Palabras clave: similitud fonética, fricativas, chino, español.

ENG Phonetic similarity between Chinese and Spanish fricatives

Abstract: In contrastive studies between languages, phonemes are typically compared to determine whether they are identical. However, only a handful of studies have examined the acoustic parameters of these phonemes to reach such an end. This paper compares and evaluates the acoustic properties of phonemes from two languages that have some similar phonemes. The languages under study are Mandarin Chinese and Spanish. Two groups of participants, native Chinese speakers and native Spanish speakers read texts containing fricatives in both languages. Comparisons were made between groups and languages. The results show that even though these two languages may share some similar fricatives, the acoustic parameters may present different values. In addition, Chinese speakers have established identical phonetic categories for both languages, whereas Spanish speakers have yet to accomplish this.

Keywords: phonetic similarity, fricatives, Chinese, Spanish.

Sumario: 1. Introducción. 2. Marco teórico. 2.1. Sistema consonántico del español y del chino. 2.2. Las fricativas del chino. 2.3. Las fricativas del español. 2.4. Objetivos del estudio. 3. Metodología. 3.1. Participantes y materiales de grabación. 3.2. Procedimiento experimental. 3.3. Tratamiento de los datos y análisis. 4. Resultados. 4.1. Diferentes grupos de hablantes utilizando una misma lengua. 4.2. El mismo grupo de hablantes utilizando lenguas diferentes. 5. Discusión de los datos. 6. Conclusiones. Agradecimientos. Contribución de autoría CREdIT. Referencias bibliográficas. Anexo.

Cómo citar: Liu, Z.; Machuca, M. J. (2025). Similitud fonética en la comparación de los fonemas fricativos del chino y del español. *Círculo de Lingüística Aplicada a la Comunicación* 103 (2025): 243-255. <https://dx.doi.org/10.5209/clac.87746>

1. Introducción

Hacer referencia a sonidos similares entre dos lenguas cuando se comparan sus sistemas fonético-fonológicos lleva a preguntarnos qué se entiende por similitud fonética. Ya Austin (1957) determinó cómo se podría definir el concepto de similitud fonética; según él, los fonemas son fonemas debido a su función y a su distribución, no a su similitud fonética. No obstante, existen muchos trabajos sobre lenguas extranjeras que intentan establecer un contraste entre el sistema fonológico de la lengua nativa y el de la lengua que se está aprendiendo a partir de la comparación de los sonidos de esas lenguas. Además, existen modelos teóricos de producción y de percepción del habla basados en la facilidad o en la dificultad de articular o percibir sonidos de acuerdo con la similitud fonética que hay entre ellos. Según Flege (2003), el grado de diferencia percibida de los sonidos de la L2 determina el modo en que los hablantes categorizarán estos sonidos, llegando a establecer incluso una nueva categoría. Si se percibe una mayor diferencia entre el sonido de la L2 y el de la L1 más próximo en cuanto a su similitud fónica, mayor será la posibilidad de que los hablantes formen una nueva categoría para el nuevo sonido. Cuando el aprendiz percibe que los sonidos de la L2 son similares a los de la L1, las categorías de la L1 y las de la L2 se fusionan formando una única categoría, aunque parece que este poder de atracción entre los sonidos de ambas lenguas es mayor si el aprendizaje de la L2 se produce en edad adulta (Baker, Trofimovich, Mack & Flege, 2002).

Las herramientas de análisis de voz permiten una investigación más precisa sobre la similitud fonética de dos o más lenguas a partir de datos sobre el número de fonemas, sobre la distribución de los alófonos o sobre las propiedades fonotácticas de esos segmentos (véase, por ejemplo, Mielke, 2012). No hay un acuerdo claro de cuál es el procedimiento más preciso para determinar la similitud fonética entre dos lenguas (Laver, 1994, p. 391). Ladefoged (1969), en uno de sus primeros trabajos, señalaba que era importante establecer el grado de similitud fonética entre segmentos para llegar a comparar dos lenguas entre sí, y añadía que los rasgos distintivos habían sido empleados como un intento de sistematizar una serie de rasgos que sirvieran para contrastar los sistemas fonológicos de lenguas diferentes. No obstante, algunos estudios han demostrado que, aun comparando la misma clase de fonemas realizados por el mismo hablante en dos lenguas diferentes, no se encuentra una similitud desde el punto de vista fonético (Liu, 2019; Sypiańska & Constantin, 2021, entre otros).

Liu (2019) compara las oclusivas sordas /p, t, k/, no aspiradas en chino, con las oclusivas sordas del español producidas por un mismo hablante; las primeras producciones son reflejo de la L1 de los participantes, las segundas corresponderían a la lengua extranjera que están aprendiendo. En principio, las oclusivas sordas no aspiradas son fonemas idénticos con relaciones fonológicas diferentes en cada una de las lenguas, pero pertenecientes a la misma clase de sonidos: en chino mandarín no existen oclusivas sonoras, pero sí aspiradas y en español no existen oclusivas sordas aspiradas, pero sí oclusivas sonoras; en ambas lenguas las oclusivas son una clase fonológica que comparten los mismos rasgos distintivos si se comparan entre ellas, aunque habría que añadir otros rasgos si se comparan con otros fonemas dentro de la misma clase. Los resultados han mostrado que existen ciertos parámetros, como la duración de la fase de silencio de las oclusivas y la del *Voice Onset Time* (VOT), que arrojan valores diferentes desde el punto de vista estadístico cuando el mismo grupo de hablantes produce una /p/ en su L1 o una /p/ en su L2. En todo caso, decidir que existe o no una similitud fonética entre dos sonidos pertenecientes a la misma clase natural en diferentes lenguas, aunque esas lenguas sean habladas por la misma persona, es arriesgado porque pueden existir otros factores externos que condicionen los datos obtenidos, como el hecho de que las duraciones, generalmente, son mayores en las lenguas que se aprenden que en la lengua nativa. Galaczi, Post, Li, Barker y Schmidt (2017) observaron que los aprendices con niveles de conocimiento más bajos producían un número de sílabas menor que los que tenían un nivel más alto; en consonancia con esto, las duraciones de las sílabas se reducían cuando los aprendices producían un mayor número de sílabas.

El objetivo de este trabajo es observar las características acústicas de los sonidos fricativos realizados por el mismo grupo de hablantes en dos lenguas diferentes, una lengua nativa y una lengua aprendida, para determinar si los aprendices crean una nueva categoría fonética diferente a la de su lengua nativa. Se comparan las realizaciones de estos fonemas en aprendices y en nativos. En el caso de los hablantes nativos del chino, se analizan esos sonidos tanto en su lengua nativa, chino mandarín, como en la lengua extranjera que aprenden, el español. En el caso de los hablantes nativos de español, se analizan estos sonidos en su lengua nativa, español, y en la lengua que aprenden, chino mandarín. Para llevar a cabo esta comparación, es necesario describir en cada una de las lenguas las características acústicas y articulatorias de los sonidos que se van a analizar.

2. Marco teórico

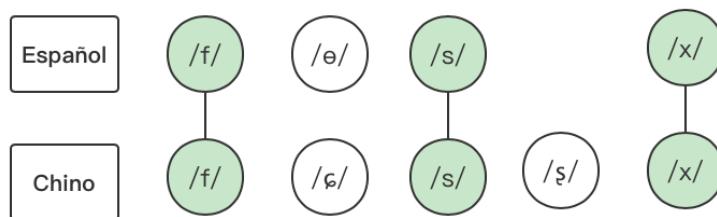
2.1. Sistema consonántico del español y del chino

Partiendo de la clasificación de los fonemas obstruyentes que establece Liu (2019) para ambas lenguas, puede observarse que las fricativas que comparten las dos lenguas son /f, s, x/ (véase Figura 1).

Se han escogido las fricativas, entre las obstruyentes de ambas lenguas, porque las oclusivas sordas, que son obstruyentes que también comparten ambas lenguas, vienen caracterizadas por un número mayor de rasgos acústicos que pueden ser diferentes en las dos lenguas analizadas debido a los rasgos propios de esta clase de sonidos: fase de silencio, fase de explosión o la duración del VOT. Las fricativas, en cambio, son consonantes que se producen con una aproximación de los órganos articulatorios que intervienen en la producción de esos sonidos, el orificio que queda entre los articuladores permite la salida de aire por la cavidad

oral, pero es tan estrecho que el aire en su salida roza contra estos órganos y se genera el ruido de fricción característico de este tipo de sonidos, por lo que en los sonidos fricativos se describen las características de esa corriente de aire que sale al exterior. Ladefoged y Maddieson (1996, p. 137) indican que estos sonidos poseen un alto grado de precisión articulatoria, y una modificación mínima de los articuladores puede provocar grandes diferencias: “one milimiter in the position of the target for the crucial part of the vocal tract makes a great deal of difference”. Esta precisión articulatoria irá asociada a unas características acústicas concretas relacionadas con la energía que se produce cuando se realizan estos sonidos en un punto de articulación concreto. Además, Ladefoged y Disner (2012, p. 168) dividen las fricativas en dos categorías, las sibilantes y las no sibilantes. Para producir un sonido sibilante, además de tener un flujo de aire que atraviesa de forma brusca los órganos articulatorios, este debe pasar por un obstáculo como los incisivos superiores. Los no sibilantes, en cambio, además de producirse mediante una constricción en algún canal estrecho en la cavidad oral, dicha constricción es menor en comparación con las sibilantes. En este trabajo, se analizarán con detalle los parámetros acústicos relacionados con la corriente del aire (energía en el espectro) y la duración de estos sonidos.

Figura 1. Comparación de las fricativas en las dos lenguas (Liu, 2019)



2.2. Las fricativas del chino

Para la clasificación de los fonemas fricativos del chino mandarín, siguiendo a la mayoría de los fonetistas, partimos de cinco fonemas fricativos sordos en la lengua china: el labiodental /f/, el dentoalveolar /s/, el retroflejo /ʂ/, el palatal /ç/ y el velar /χ/ (Bao & Lin, 2014; Duanmu, 2007; Lee, Zhang, Li, Tao & Bond, 2012) que se representan respectivamente por *f*, *s*, *sh*, *x*, *h* en el sistema romanizado (*pinyin*).

La consonante /f/ se realiza en chino como labiodental sorda, se articula acercando el labio inferior a los incisivos superiores formando un canal estrecho por el que pasa el aire. En cuanto a los sonidos fricativos que se realizan en la zona alveolopalatal, es decir, /s/, /ʂ/ y /ç/, existe una controversia sobre los puntos de articulación. Lu (2014) indica que, aunque siempre se ha dicho que el fonema /ç/ se presenta en distribución complementaria con el fonema /s/, pues el primer sonido siempre se encuentra ante vocales anteriores, por lo que tendría una realización más cercana al paladar, y el segundo nunca aparece ante este tipo de vocales (una realización más alveolar), los resultados de un test de percepción en el que los sujetos debían identificar un continuo de un sonido a otro sugieren que la distribución no fuerza a los hablantes del chino a analizar estos dos sonidos como variantes de una misma categoría. En este sentido, existen autores que ya habían indicado que en la región alveolar y en la posalveolar se produce un contraste entre las tres fricativas sibilantes sordas en chino mandarín /s, ʂ, ç/ (Duanmu, 2007). Ladefoged y Maddieson (1996) han señalado también que, en la realización de estos tres sonidos, los incisivos superiores y los inferiores se encuentran muy próximos, rasgo que indica que son fricativas claramente sibilantes. En los cortes sagitales que se muestran en este mismo manual se indica que, en el caso de la /s/, el punto de constricción es un poco distinto en cada hablante, puede variar desde la parte trasera de los incisivos superiores hasta la zona alveolar (Ladefoged & Maddieson, 1996, p. 151), aunque Stokes y Zhen (1998) señalan en un estudio electropalatográfico que es alveolar. En cambio, la /ʂ/ presenta, según estos autores, un gesto articulatorio muy diferente a lo que se considera un sonido retroflejo, pues este sonido no se realiza con el ápice de la lengua elevado y curvado hacia atrás, sino que se realiza con la superficie de la lengua plana, haciendo que este sonido sea laminal más que apicoalveolar. Por último, el fonema velar /χ/ tiene un punto de articulación que puede depender de la vocal que le sigue y es la fricativa que presenta mayor variación en chino (Bao & Lin, 2014). Según Trísková (2008), este fonema puede presentar tres realizaciones posibles: velar, uvular y glotal, por lo que no presenta un único punto de articulación.

Si consideramos las características acústicas de estos sonidos, Lee, Zhang y Li (2014) han analizado las fricativas del chino atendiendo a los parámetros temporales, espectrales y frecuenciales. Según estos autores, la zona en la que se produce el contacto entre los articuladores y el hecho de que sea sibilante o no determinan las características de la energía spectral. El pico máximo varía en función del punto de articulación para las sibilantes, cuanto más posterior sea el punto de articulación, menor valor de frecuencia alcanza el pico máximo, ya que una constricción más posterior se asocia con zonas de baja frecuencia. Zhang (2010) señala que los parámetros espectrales reflejan, en general, la forma de un sonido: cuanto más posterior sea el lugar de constricción en la cavidad oral, más claras son las características de la energía en el espectro. Lo mismo sucede en el espectro de una fricativa sibilante, que suele presentar una forma bien definida, frente a las no sibilantes, que muestran una forma plana. Por lo tanto, según estos autores, la posterioridad de una fricativa indica que el pico máximo de energía se situará en una zona de frecuencia más baja,

y la asibilación favorece que el espectro posea una envolvente bien definida. Siguiendo a estos autores, la distribución frecuencial de la energía de /f/ es dispersa, se extiende desde 8000 Hz a 12000 Hz. La zona de concentración de energía y el límite inferior de fricción son altos, normalmente por encima de 5000 Hz. La zona de concentración para la /s/ se encuentra en una zona de frecuencia alta, y el límite inferior de la zona de fricción se puede apreciar de una forma clara, ya que también se trata de una de las características de los sonidos fricativos cuyo punto de articulación se encuentra entre los alveolos y el paladar. La fricativa velar /x/, por su parte, no presenta rasgos específicos que la caractericen y estos pueden variar en función de la vocal siguiente. Como la energía de esta fricativa generalmente es débil y su duración es breve, si se encuentra en posición interior de palabra, a veces no se percibe esta fricción; es como si fuera un apoyo consonántico entre las dos vocales (Bao & Lin, 2014).

2.3. Las fricativas del español

Al igual que sucedía en chino, no existe un acuerdo sobre el número de fricativas en español. Según Quilis y Fernández (1985, p. 91), el español tiene cinco fonemas fricativos, los cuatro primeros sordos y el último sonoro: el labiodental /f/, el interdental /θ/, el alveolar /s/, el velar /x/ y el palatal /j/. No obstante, este número se reduce a cuatro según Fernández Planas (2005), prescindiendo de la palatal /j/. Machuca (2000) señala que la realización de /j/ es variable y presenta realizaciones oclusivas o africadas en posición inicial y detrás de nasal y lateral, y aproximantes en los otros contextos, además de otras realizaciones fricativas que pueden encontrarse en el español de América. El hecho de que para el fonema /j/ existan diferentes realizaciones debidas tanto a factores individuales del hablante como a factores geográficos hace que en este trabajo partamos de la clasificación de Fernández Planas (2005), en la que se considera que el español tiene cuatro fonemas fricativos, todos ellos sordos.

La fricativa /f/ es labiodental sorda, aunque en los contextos en los que se encuentra seguida de una consonante sonora se da una realización sonorizada (Real Academia Española – Asociación de Academias de la Lengua Española, 2011). La fricativa interdental sorda /θ/ se realiza colocando el ápice de la lengua entre los incisivos superiores e inferiores; también puede sonorizarse. La /s/ es alveolar sorda, a pesar de que también posee unos alófonos según las diferencias regionales (véase, por ejemplo, Campos-Astorkiza, 2019; Martínez Celrá, 1996). Además, también es posible el alófono sonoro de este mismo fonema cuando va seguido de una consonante sonora (Fernández Planas, 2005; Martínez Celrá, 1984; Quilis & Fernández, 1985). La fricativa velar sorda /x/ presenta la constrección en la zona velar con la aproximación del posdorso de la lengua hacia el velo del paladar. En cuanto a su punto de articulación, existen autores (Fernández Planas, 2005; Hidalgo & Quilis Merín, 2004; Martínez Celrá, 1984) que señalan que este puede retraerse hacia la zona uvular, llegando a unas realizaciones velofaríngeas, si le siguen vocales posteriores. Además, también existen realizaciones glotales en ciertas zonas de Hispanoamérica y de España.

Si nos centramos en las características acústicas de estos sonidos, podemos señalar que la energía de la fricativa labiodental /f/ tiene una intensidad baja (Martínez Celrá, 1984), aunque suele extenderse, al igual que ocurre con la consonante interdental /θ/, por todo el espectrograma, hasta llegar a los 8000 Hz (Real Academia Española – Asociación de Academias de la Lengua Española, 2011), presentando valores de frecuencia más bajos que los observados para el chino mandarín. No obstante, Barreiro Bilbao (1999), que ha analizado las fricativas no sibilantes /f/ y /θ/ del español en posición aislada, atendiendo a los picos máximos y a la frecuencia en la que se encuentran, sugiere que los componentes de frecuencia por encima de los 8000 Hz aportan una información acústica importante para la interdental /θ/, mientras que, para la labiodental, la información relevante se encuentra por debajo de los 6000 Hz, y los componentes de frecuencias por encima de este valor no son muy útiles.

En cuanto a la realización de la fricativa alveolar /s/, presenta una zona de ruido de gran intensidad y se sitúa por encima de los 4000 Hz. Su variante apical muestra una energía que llega a unos 8000 Hz y la variante predorsal, por encima de este valor. El alófono predorsal y el apical tienen manifestaciones acústicas diferentes; el predorsal tiene la energía a partir de los 5000 Hz y el apical, a partir de los 4000 Hz (Real Academia Española – Asociación de Academias de la Lengua Española, 2011).

La fricativa /x/ se caracteriza por la presencia de más de una zona de concentración de energía, que representa variaciones en la intensidad (Real Academia Española – Asociación de Academias de la Lengua Española, 2011), y, además, es diferente en función de la vocal que le siga.

En resumen, podríamos decir que aunque todas las fricativas se observan como inarmónicas en el espectrograma, hay algunas diferencias entre ellas (Fernández Planas, 2005, p. 145). Si se ordenan en función de la intensidad, la alveolar /s/ es la más intensa, seguida de la /x/, /θ/ y /f/. La zona de frecuencia en la que empieza la fricción es más baja en /x/, seguida de /f/ y /θ/, llegando al punto máximo en /s/. Si consideramos la zona de frecuencia de mayor intensidad de cada una de las cuatro fricativas, la velar /x/ presenta la zona de frecuencia de mayor intensidad en la zona más baja, a unos 1700 Hz; la /s/ la tiene a 3500 Hz; la /f/, a 5000 Hz; y la /θ/, a 7500 Hz. Cuando las fricativas están sonorizadas, se puede observar, además de las características de sus alófonos sordos, la presencia de un formante de sonoridad en zonas de baja frecuencia.

La Tabla 1 muestra los valores medios del pico máximo de las tres fricativas analizadas en Manrique y Massone (1981) producidas por cuatro hablantes masculinos de Argentina en dos condiciones: cuando la fricativa se produce de forma sostenida o cuando se combina con cada una de las cinco vocales.

Tabla 1. Valores del pico máximo de las fricativas del español de Argentina presentados en el estudio de Manrique y Massone (1981)

Fonema	Sostenida	[i]	[e]	[a]	[o]	[u]
/f/	1733	2150	1700	1633	1633	1466
/s/	5300	5560	5160	5350	4333	4500
/x/	1462	2950	2075	1600	1025	900

Antes de acabar esta descripción, es necesario detenerse en los resultados obtenidos en el estudio de Cicres (2011), pues ha considerado para el español no solo el pico máximo de energía, como aparece también en los trabajos mencionados, sino también cuatro parámetros más relacionados con el espectro, que también se analizarán en este trabajo: el centro de gravedad (*Center of gravity*, de aquí en adelante en sus siglas en inglés, COG), que cuantifica la concentración de energía promedio; la desviación estándar (*Standard deviation*, SD), que mide la distancia de la distribución de frecuencias desde la frecuencia media; la asimetría, que está vinculada con la distribución de energía a los lados del centro de gravedad –una asimetría con valor 0 implica una forma equilibrada, mientras que un valor positivo señala una mayor concentración de energía en zonas de energía de baja frecuencia y un valor negativo, en zonas de energía de alta frecuencia–; la curtosis, que especifica el grado de apuntamiento de la forma espectral, –una curtosis positiva indica una forma puntañada, mientras que una curtosis negativa implica una forma plana–.

Los resultados obtenidos en su experimento (véase Tabla 2) muestran que los valores del pico máximo de frecuencia para la fricativa velar /x/ son más bajos que los de las fricativas labiodental y alveolar. Sin embargo, la desviación estándar de esta variable es menor para las fricativas alveolar y velar que para las otras fricativas. El centro de gravedad es menor para la /x/ que para las otras fricativas, aunque la desviación estándar de este parámetro es menor tanto para la fricativa velar como para la alveolar. La desviación estándar tiene valores más bajos en la región alveolar y valores más elevados en las regiones labiodental e interdental. Por último, la curtosis y la asimetría en la dental proporcionan valores más bajos que los de las otras fricativas, especialmente si la comparamos con la velar.

Tabla 2. Valores extraídos de los parámetros espectrales de las fricativas del español, según el estudio de Cicres (2011). COG es el centro de gravedad y la SD, la desviación típica.

Fonema	Pico máximo	COG	SD	Curtosis	Asimetría
/f/	4767,38 (2144.05)	6198,24 (1810.41)	2774,87 (624.08)	1,08 (3.23)	0,56 (0.94)
/θ/	3474,88 (1405.72)	6167,30 (1032.46)	3789,38 (481.22)	0,31 (1.80)	0,49 (0.57)
/s/	4424,20 (364.20)	5231,07 (414.14)	1932,08 (324.48)	3,89 (2.57)	1,08 (0.55)
/x/	1748,16 (321.04)	2829,65 (292.01)	2269,56 (271.96)	8,39 (4.05)	2,48 (0.50)

2.4. Objetivos del estudio

El objetivo principal de este trabajo es comparar fonemas comunes en dos lenguas para determinar si esa similitud fonológica viene respaldada por las características acústicas de los sonidos en cada una de las lenguas analizadas. El chino y el español, aunque proceden de familias lingüísticas muy diferentes, pueden contener dentro de sus inventarios fonológicos fonemas comunes. En el caso de las fricativas, el chino mandarín comparte con el español las consonantes /f, s/. En ambas lenguas el primero es no sibilante y el segundo, sibilante. Se podría pensar que el fonema /x/ también es un fonema compartido en ambas lenguas porque el símbolo empleado siguiendo la transcripción del Alfabeto Fonético Internacional es el mismo, no obstante, como ya se ha mencionado, es el fonema que sufre mayor variación en chino produciéndose, incluso, variaciones dentro del mismo hablante (Trísková, 2008, p. 4). La mayoría de los investigadores consideran esta consonante velar (Ladefoged & Wu, 1984; Lin, 2007), otros la describen con una realización uvular (Pulleyblank, 1984) y otros afirman que se puede realizar como glotal (Duanmu, 2007), aunque esta última realización no se considere como estándar. En español, ocurre algo similar, la realización del fonema fricativo velar es generalmente velar, pero puede retraerse hacia la zona uvular o realizarse como velofaríngea ante un contexto vocálico velar (Fernández Planas, 2005; Hidalgo & Quilis Merín, 2004; Martínez Celdrán, 1984, entre otros). Además, también se pueden encontrar realizaciones glotalizadas en algunas zonas del sur de España y de Hispanoamérica. Por esta razón, los fonemas fricativos considerados comunes en nuestro estudio se reducirán al labiodental y al alveolar.

Las realizaciones de los dos fonemas seleccionados, producidas por un grupo de hablantes nativos de chino mandarín con un nivel de conocimiento alto en español, se compararán en las dos lenguas y también se compararán con realizaciones de hablantes nativos de español. El mismo proceso se llevará a cabo con los hablantes nativos de español que tienen un nivel alto de chino mandarín. Primero, se compararán estos fonemas en las dos lenguas y, después, se comparará la realización de los fonemas chinos de los hablantes aprendices con realizaciones de hablantes nativos de chino mandarín. Las preguntas que pretendemos contestar en este estudio son:

- Si los hablantes de diferentes grupos (nativos y aprendices) presentan las mismas características acústicas cuando consideramos una misma lengua.
- Si los hablantes de un mismo grupo presentan los mismos valores para las dos lenguas analizadas.

3. Metodología

3.1. Participantes y materiales de grabación

Para este tipo de estudios se suele utilizar lectura de palabras aisladas como método de obtención de datos, pero, en este trabajo, se ha decidido emplear la lectura de textos y evitar así la entonación propia de un listado de palabras.

El corpus utilizado para recoger las muestras en chino procede de dos textos extraídos de Liu (2019). Uno de los textos era fonéticamente equilibrado y otro texto, aunque no estaba fonéticamente equilibrado, servía para conseguir más casos de las consonantes en diferentes contextos vocálicos. Para las muestras en la lengua española, se emplean dos textos fonéticamente equilibrados (Bruyninckx et al., 1994; Ortega-García et al., 2000) que ya se han utilizado ampliamente en la investigación fonética. El de Bruyninckx et al. (1994) se diseña para investigar la influencia de las lenguas sobre la calidad de voz de los informantes, y el de Ortega-García et al. (2000) forma parte del corpus AHUMADA y se diseñó con fines judiciales.

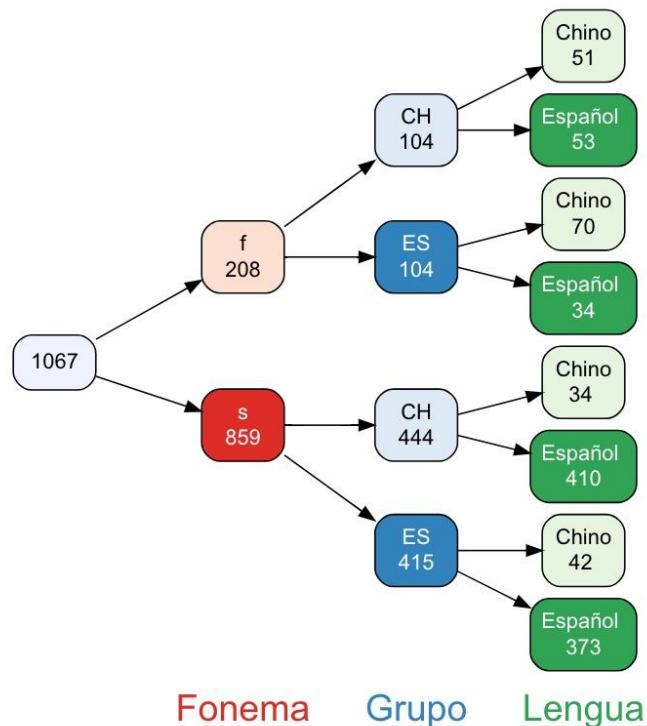
En este estudio han participado un total de ocho informantes, todas femeninas, cuyo rango de edad se sitúa entre 21 y 25 años. Se dividen en dos grupos en función de su perfil lingüístico. El primer grupo consta de cuatro hablantes nativos del chino mandarín, que estaban cursando en el momento de las grabaciones el máster de Lengua Española, Literatura Hispánica y Español como Lengua Extranjera en la Universidad Autónoma de Barcelona. Todas ellas habían cursado la carrera de Filología Española y estaban aprendiendo el español como lengua extranjera. El otro grupo está formado por cuatro hablantes nativas del español, que estaban cursando el grado de Traducción e Interpretación (lengua chino-español) en la misma institución y que estaban aprendiendo chino como lengua extranjera. Ambos grupos comparten la similitud de haber adquirido el inglés como una segunda lengua. Además, todos poseen un nivel intermedio (B2/C1) para su L2 según el Marco Común Europeo de Referencia para las Lenguas.

3.2. Procedimiento experimental

Los hablantes leyeron estos textos a un ritmo que consideraban adecuado en una sala insonorizada del Servei de Tractament de la Parla i el So (STPS) de la UAB. Antes de la lectura, se mantuvo una charla de unos diez minutos con las informantes en la lengua que iban a grabar para facilitar la inmersión en esa lengua. Las grabaciones tuvieron lugar en distintos días para cada lengua con el fin de evitar al máximo posible la influencia de la otra lengua.

Para llevar a cabo la grabación, se utilizó el micrófono Rode NT1-A, con una respuesta de frecuencia entre 20 y 20000 hercios. Su rango dinámico era de 132 dB. Todas las grabaciones se obtuvieron en el formato de audio digital y fueron digitalizadas a una frecuencia de muestreo de 44 kHz.

Figura 2. Porcentaje y número de casos analizados en este estudio según el fonema, el grupo y la lengua



Las muestras de habla fueron segmentadas de manera diferente para cada lengua en la herramienta Praat (Boersma & Weenink, 2017). Para los textos de español, se utilizó la segmentación automática con la herramienta EasyAlign (Goldman & Schwab, 2014) y una posterior revisión manual de esa segmentación; para los textos en chino, se procedió a una segmentación manual debido a que EasyAlign no ofrecía segmentación automática para la lengua china, pero se siguieron los mismos criterios de segmentación que se encontraban en la herramienta. Posteriormente, se etiquetaron, a partir de las características identificadas en los espectrogramas, aquellas realizaciones consideradas correctas; estas realizaciones son las que se analizaron en este estudio. Por ejemplo, se excluyeron de este análisis aquellos casos de sonorización de estos fonemas o realizaciones oclusivas de estas fricativas, entre otros. Además, puesto que las reglas fonotácticas del chino no permiten que estos fonemas fricativos se encuentren en la posición de coda, en este estudio solo se consideraron aquellas realizaciones que se encuentran en el ataque silábico.

De esta manera, el corpus analizado en este estudio consta de 548 casos en el grupo chino y 519 en el grupo español. La Figura 2 presenta el número de casos analizados para cada lengua, cada grupo y cada fonema. La diferencia entre el número de casos del fonema /s/ en chino y en español se debe a la frecuencia de aparición de estos fonemas en cada una de las lenguas (véase Machuca & Ríos, 2017 para unos datos más exhaustivos sobre la frecuencia de aparición de este fonema en español).

3.3. Tratamiento de los datos y análisis

El análisis estadístico y la visualización de los datos se han realizado mediante el programa RStudio (RStudio Team, 2022). Para la comparación de los datos, se ha llevado a cabo la prueba de *t* utilizando el paquete *ggstatsplot* (Patil, 2021). La variable dependiente está formada por cada uno de los parámetros acústicos y la variable independiente, por el grupo o la lengua teniendo en cuenta la pregunta a la que se está contestando.

Los parámetros acústicos que se analizan para estos fonemas se pueden dividir en parámetros temporales, la duración del segmento, y parámetros espectrales, que incluyen los cinco parámetros relacionados con la forma espectral que presentan las consonantes, es decir, el pico máximo de energía y los cuatro momentos espectrales que ya habíamos descrito (Cicres, 2011). Para las pruebas estadísticas, salvo para la asimetría y la curtosis, por presentar valores negativos, se ha partido de los datos en escala logarítmica como modo de normalización entre los hablantes. Pese a eso, para una mejor comprensión, se presentan los valores numéricos en las tablas ilustrativas y en las figuras.

4. Resultados

En la Tabla 3, se pueden observar los valores medios de los diferentes parámetros acústicos considerados en este estudio. Como se aprecia en esta tabla, la duración es el parámetro que presenta unos resultados más sólidos, ya que, en la mayoría de los casos, las consonantes producidas por un aprendiz duran más que las realizadas por un hablante nativo. Los demás parámetros presentan más variabilidad, como veremos a continuación.

Tabla 3. Valores medios de los diferentes parámetros en función del fonema, grupo y lengua.
Int.media hace referencia a la intensidad media; *SD*, a la desviación estándar; *COG*, al centro de gravedad.

Fonema	Grupo	Lengua	Duración	Pico máximo	Int. media	SD	Asimetría	Curtosis	COG
/f/	CH	Chino	102,69	6631,45	58,27	4567,37	0,10	0,31	6991,35
	CH	Español	108,19	6083,13	57,64	4153,45	0,73	2,04	5911,75
	ES	Chino	133,01	2487,14	52,90	3545,74	1,68	2,76	3432,76
	ES	Español	90,91	3797,76	55,71	4143,53	0,83	0,30	5106,79
/s/	CH	Chino	123,56	8232,76	66,03	2991,54	-0,80	3,43	8453,49
	CH	Español	105,19	7622,36	68,78	2917,54	-0,46	1,81	7619,25
	ES	Chino	161,36	7145,76	63,88	2537,26	0,05	1,61	7024,47
	ES	Español	86,90	7405,23	65,81	2870,37	-0,31	0,99	6715,11

4.1. Diferentes grupos de hablantes utilizando una misma lengua

Si tenemos en cuenta los valores para el fonema labiodental en el chino mandarín (Tabla 3), considerando que para un grupo constituye la lengua nativa (CH_chino) y para el otro, la aprendida (ES_chino), los valores son mayores para el nativo que para el aprendiz en todos los casos, excepto en la duración de la consonante, la asimetría y la curtosis. No obstante, si seleccionamos como lengua el español, los hablantes nativos de español (ES_español) presentan todos los valores más bajos que el aprendiz (CH_español), excepto para la asimetría.

En el caso del fonema alveolar (Tabla 3), si comparamos aprendices (ES_chino) y nativos (CH_chino) hablando chino mandarín, se observan de nuevo valores más altos en los nativos chinos que en los aprendices, excepto para la duración de la consonante y para la asimetría. En cambio, si nos fijamos en el español (CH_español, aprendices, frente a ES_español, nativos), al igual que hemos visto para la labiodental, los hablantes nativos de español muestran todos los valores más bajos que los aprendices, excepto para la asimetría.

En cuanto a los valores de significación de las variables analizadas (Tabla 4 del Anexo), para el fonema /f/, se observa una diferencia significativa debida al grupo en todos los parámetros analizados en la lengua china. Como ya se ha mencionado, en general, los valores para la duración, la asimetría y la curtosis son más altos en el grupo de aprendices, que en el grupo de los hablantes nativos cuando se considera el chino mandarín. No obstante, si observamos este mismo fonema en la lengua española, casi no se muestran diferencias significativas entre los dos grupos, salvo para la duración, el pico máximo y la intensidad media, que presentan unos valores más altos en el grupo de aprendiz.

En cuanto al fonema /s/, a diferencia de /f/, casi no se observa una diferencia significativa entre los dos grupos para la lengua china, salvo para la variable duración, la asimetría y el centro de gravedad. Para la lengua española, existen más variables que muestran la diferencia entre producir ese fonema como nativo del español o como aprendiz (véase Tabla 4 en el Anexo).

4.2. El mismo grupo de hablantes utilizando lenguas diferentes

Si tomamos al mismo grupo de hablantes produciendo un fonema fricativo labiodental en su lengua nativa (CH_chino, ES_español) y en la lengua que está adquiriendo (CH_español, ES_chino), en el caso del grupo de hablantes chinos podemos observar que, aunque presentan valores más bajos cuando hablan español, excepto para la duración, la curtosis y la asimetría, la diferencia entre esos valores es menor que cuando se comparaban diferentes grupos de hablantes. Para el grupo de hablantes españoles (ES_español y ES_chino) sucede lo mismo, cuando hablan en chino, la lengua que aprenden, los valores son más bajos que cuando hablan en su lengua nativa, excepto para la duración, la curtosis y la asimetría.

En la producción del fonema alveolar por parte del grupo de hablantes nativos del chino cuando habla en chino (CH_chino), ese fonema arroja valores más altos que cuando habla en la lengua que aprenden (CH_español), excepto en los valores de intensidad. Sin embargo, en el caso de que los hablantes pertenezcan al grupo de hablantes españoles, cuando hablan en chino (ES_chino), la lengua que aprenden, mantienen valores más elevados en la duración, en la asimetría y en la curtosis, pero también en el centro de gravedad que cuando hablan en su lengua nativa (ES_español).

En cuanto a las diferencias estadísticas debidas a las lenguas en cada grupo (véase el Anexo), para el fonema /f/, no se observa una diferencia significativa en el grupo cuya lengua materna es el chino. Para el fonema /s/, solo hay una diferencia en la variable duración, que presenta unos valores más altos en la lengua nativa de los hablantes (chino) que en la lengua que aprenden.

Si consideramos el grupo cuya lengua materna es el español, se aprecia que muchas variables presentan valores significativamente diferentes para cada lengua. Si observamos el fonema /f/, solo la variable del pico máximo no ha arrojado una diferencia significativa. Con respecto al fonema /s/, la duración, la asimetría, la curtosis y el centro de gravedad arrojan unos valores más altos en la lengua que aprenden los hablantes que en su lengua nativa.

5. Discusión de los datos

Si observamos los valores acústicos que se han obtenido en cada caso, independientemente de las lenguas, el pico máximo siempre es más alto en el fonema sibilante, /s/, que en el no sibilante, /f/. Esta observación es contraria a lo que, en principio, aseguran en los diferentes trabajos, en los que se indica que el pico de energía para la /s/ es menor que el de la /f/, pues hay una relación entre la zona en la que está situado el pico y el punto de articulación de la consonante: cuanto más posterior sea el punto de articulación, el parámetro de mayor intensidad de la energía se encontrará en una zona de frecuencia más baja. En este sentido, existen estudios que señalan (C. Y. Lee et al., 2014) que el fonema /s/ en chino presenta mayores valores para el pico máximo que los otros fonemas, hecho que se muestra en nuestros datos tanto para el chino como para el español. Quizá se debería pensar que este incremento en la localización del pico de energía de /s/ respecto al de /f/ esté más relacionado con el rasgo sibilante; como ya se ha mencionado la /s/ es sibilante y la /f/ no lo es.

Los resultados también muestran que las lenguas poseen sus características propias, es decir, si nos fijamos en los valores del pico máximo de la labiodental, por ejemplo, podemos observar que en su lengua nativa los hablantes nativos del chino siempre presentan un valor más alto (6631,45 Hz) que los hablantes nativos de español (3797,76 Hz). Esta característica se da generalmente en los parámetros analizados, lo que trae consigo que cuando un hablante nativo de chino realiza una labiodental en español siempre presenta valores de pico más altos que los propios nativos de esa lengua, porque intrínsecamente esa lengua posee valores más altos del pico espectral. Si consideramos que el valor del pico espectral y el centro de gravedad se corresponden con la zona de frecuencia en la que hay mayor concentración de energía en el espectro y que existe, como aseguraban Jongman, Wayland y Wong (2000), una relación entre el lugar de articulación y el pico espectral en el sentido de que el valor del pico disminuye cuanto más retrasado sea el punto de articulación de la consonante, quizás estos resultados nos indiquen una realización más anterior de este fonema para los hablantes chinos que para los hablantes españoles. También podemos relacionarlo con la longitud

del tracto vocal que habían señalado Hughes y Halle (1956): "we expect an inverse relationship to hold between the length of the effective portion of the vocal tract and the frequency of the peak". Los resultados mostrarían que la longitud del tracto cuando se realiza este fonema sería mayor en el caso de los hablantes españoles, pues presentan valores del pico mucho más bajos.

Lo mismo ocurre con el fonema alveolar sibilante; el chino se caracteriza por presentar un pico espectral (8232,76 Hz) y un centro de gravedad (8453,49 Hz) en una región de frecuencia más alta que el español (7405,23 y 6715,11 Hz, respectivamente), por lo que cuando un chino realiza el fonema /s/ en español, esa frecuencia sigue estando localizada en una zona de frecuencia más alta que un hablante nativo español. Por esta razón, cuando consideramos las lenguas, el hablante chino aprendiz de español (CH_español), aunque presente valores más bajos de frecuencia en la energía espectral que cuando la realiza en su lengua nativa, siempre presentará valores más altos que un hablante nativo de la lengua que está aprendiendo, español, y viceversa, cuando el español aprendiz de chino (ES_chino) realiza un fonema fricativo, presenta la energía espectral en una zona de frecuencia más alta que cuando lo produce en su lengua nativa, pero no llega a los valores de la región de frecuencia del hablante nativo del chino.

Podríamos señalar que en la similitud fonética de los sonidos de dos lenguas diferentes, las características acústicas intrínsecas de cada lengua influyen sobre las características de la lengua que se aprende. Puede que estos fonemas sean comunes a ambas lenguas si nos fijamos en su sistema fonológico, pero no si nos fijamos en su sistema fonético. Las hablantes chinas, a la vista de los resultados obtenidos, ya han igualado ambas categorías fonéticas, por eso apenas hay diferencias significativas entre ellas, mientras que las españolas, quizá por un nivel de conocimiento más bajo, todavía intentan realizar categorías fonéticas diferentes, como se muestra en las diferencias significativas observadas tanto para /f/ como para /s/.

Por último, los valores de duración de las consonantes muestran que los hablantes aprendices presentan una duración mayor de los fonemas cuando pertenecen a una lengua extranjera que en su propia lengua, a pesar de que los fonemas son comunes a los dos sistemas fonológicos. Galaczi et al. (2017) ya habían señalado que los hablantes extranjeros tienden a realizar duraciones silábicas más largas, y que se van reduciendo en función del conocimiento de la lengua extranjera.

6. Conclusiones

Si atendemos a las preguntas que nos hemos planteado en este trabajo, podemos señalar que cuando consideramos una misma lengua observamos diferencias entre el chino y el español, aunque los fonemas analizados (/s/ y /f/) sean comunes en ambas lenguas. El chino presenta en los dos fonemas características de la energía en una zona de mayor frecuencia que el español, por lo que, cuando aprenden una lengua extranjera como el español, realizan esos fonemas también con una energía localizada en una región de frecuencia más alta que los hablantes nativos de esa lengua, mientras que en español esa energía espectral está localizada en una frecuencia más baja, por lo que, cuando realizan esos mismos fonemas en una lengua extranjera como el chino, la zona en la que la fricción es más intensa se sitúa en una zona más baja que la de los hablantes nativos de esa lengua. Estos resultados nos ayudan a afirmar que existe una influencia de la lengua nativa que se manifiesta en la energía espectral de estos fonemas fricativos comunes.

En cuanto a la pregunta sobre si los hablantes de un mismo grupo presentan los mismos valores para las dos lenguas analizadas, podemos afirmar que, en el caso de los hablantes chinos, tanto si hablan chino, su lengua nativa, como si hablan español, su lengua extranjera, los resultados, en general, no difieren entre sí. Eso quiere decir que los fonemas de las dos lenguas pertenecen a la misma categoría fonética. Sin embargo, en el caso de los hablantes españoles, los datos difieren en gran medida cuando estos mismos hablantes producen estos fonemas en su lengua nativa o en su lengua extranjera, como lo indican las diferencias significativas encontradas en todos los parámetros, excepto para el pico máximo. Por tanto, estos hablantes aún no han igualado esas categorías fonéticas; quizás esto nos esté mostrando que el nivel de conocimiento de chino por parte de los hablantes españoles es más bajo que el nivel de español de los hablantes chinos, aunque la acreditación que tengan sobre el conocimiento de la lengua extranjera sea similar.

Así pues, las categorías fonológicas de una determinada lengua pueden ser similares a otra, pero las características acústicas pueden ser intrínsecas de cada lengua, como se ha observado aquí con los fonemas /s/ y /f/, cuya fricción está localizada en una región de frecuencia mayor en chino que en español, indicando que las categorías fonéticas son diferentes. Se necesitaría, de todas formas, un estudio más exhaustivo en cada una de las lenguas analizadas para determinar por qué la fricativa labiodental muestra en ambas lenguas una zona de fricción localizada en una región de frecuencia más baja que la alveolar, contrariamente a lo que se indica en los estudios consultados.

Agradecimientos

Este trabajo ha recibido el apoyo del Centro de Investigación sobre la Enseñanza de Idiomas Clave de la Escuela de Lenguas Extranjeras, así como una financiación por parte de la Escuela de Estudios de Posgrado de la Universidad Normal del Este (número de proyecto: PX-49252410).

Contribución de autoría CREDIT

Las dos autoras han participado en igual medida en cada parte de la investigación y en la redacción, revisión y edición de este artículo. Ambas autoras han leído y aceptado la versión publicada del manuscrito.

Referencias bibliográficas

- Austin, William M. (1957). Criteria for phonetic similarity. *Language*, 33(4), 538-544. <https://doi.org/10.2307/411308>
- Baker, Wendy, Trofimovich, Pavel, Mack, Molly, & Flege, James E. (2002). The effect of perceived phonetic similarity on non-native sound learning by children and adults. En A. Do, L. Dominguez, & A. Johansen (Eds.), *BUCLD 26: Proceedings of the 26th annual Boston University Conference on Language Development* (p. 12). Cascadilla Press.
- Bao, Huaiqiao, & Lin, Maocan. (2014). *Shiyan Yuyinxue Gaiyao. Fonética experimental (edición revisada)*. Peking University Press.
- Barreiro Bilbao, Silvia C. (1999). Análisis acústico comparado de las fricativas castellanas no sibilantes en realizaciones aisladas. *Contextos*, 33-36, 243-261.
- Boersma, Paul, & Weenink, David. (2017). *Praat: Doing phonetics by computer*, versió 6.0.36. 5. <http://www.praat.org>
- Bruyninckx, Marielle, Harmegnies, Bernard, Llisterri, Joaquim, & Poch-Olivé, Dolors. (1994). Language-induced voice quality variability in. *Journal of Phonetics*, 22, 19-31. [https://doi.org/10.1016/S0095-4470\(19\)30265-7](https://doi.org/10.1016/S0095-4470(19)30265-7)
- Campos-Astorkiza, Rebeka. (2019). Modeling assimilation: The case of sibilant voicing in Spanish. En Mark Gibson & Juana Gil (Eds.), *Romance phonetics and phonology* (pp. 241-275). Oxford University Press.
- Cicres, Jordi. (2011). Los sonidos fricativos sordos y sus implicaciones forenses. *Estudios filológicos*, 48, 33-48. <https://doi.org/10.4067/S0071-17132011000200003>
- Duanmu, San. (2007). *The Phonology of Standard Chinese* (2nd ed). Oxford University Press.
- Fernández Planas, Ana María. (2005). Así se habla. *Nociones fundamentales de fonética general y española. Apuntes de catalán, gallego y euskera*. Horsori.
- Flege, James Emil. (2003). Assessing constraints on second-language segmental production and perception. *Phonetics and phonology in language comprehension and production: Differences and similarities*, 6, 319-355.
- Galaczi, Evelina, Post, Brechtje, Li, Aike, Barker, Fiona, & Schmidt, Elaine. (2017). Assessing Second Language Pronunciation: Distinguishing Features of Rhythm in Learner Speech at Different Proficiency Levels. En Talia Isaacs & Pavel Trofimovich (Eds.), *Second Language Pronunciation Assessment* (Vol. 107, pp. 157-182). Multilingual Matters / Channel View Publications. <https://www.jstor.org/stable/10.21832/j.ctt1xp3wcc.13>
- Goldman, Jean-Philippe, & Schwab, Sandra. (2014). EasyAlign Spanish: An (semi-) automatic segmentation tool under Praat. *Fonética experimental, educación superior e investigación*, 629-640. <http://latlcui.unige.ch/phonetique/easyalign/GoldmanSchwab-EasyAlignSpanish-5thCFE-2011.pdf>
- Hidalgo, Antonio, & Quilis Merín, Mercedes. (2004). *Fonética y fonología españolas*. Tirant Lo Blanch.
- Hughes, George W., & Halle, Morris. (1956). Spectral Properties of Fricative Consonants. *The Journal of the Acoustical Society of America*, 28(2), 303-310. <https://doi.org/10.1121/1.1908271>
- Jongman, Allard, Wayland, Ratree, & Wong, Serena. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America*, 108(3), 1252. <https://doi.org/10.1121/1.1288413>
- Ladefoged, Peter. (1969). The Measurement of Phonetic Similarity. *International Conference on Computational Linguistics COLING 1969: Preprint No. 57*. COLING 1969, Sånga Säby, Sweden. <https://doi.org/10.3115/990403.990460>
- Ladefoged, Peter, & Disner, Sandra Ferrari. (2012). *Vowels and consonants* (3rd ed). Wiley-Blackwell.
- Ladefoged, Peter, & Maddieson, Ian. (1996). *The sounds of the world's languages*. Blackwell Publishers.
- Ladefoged, Peter, & Wu, Zongji. (1984). Place of articulation: An investigation of Pekingese fricatives and affricates. *Journal of Phonetics*, 12, 267-278. [https://doi.org/10.1016/S0095-4470\(19\)30883-6](https://doi.org/10.1016/S0095-4470(19)30883-6)
- Laver, John. (1994). *Principles of phonetics*. Cambridge University Press.
- Lee, Chao-Yang, Zhang, Yu, & Li, Ximing. (2014). Acoustic characteristics of voiceless fricatives in Mandarin Chinese. *Journal of Chinese Linguistics*, 42(1), 150-171.
- Lee, Chao-Yang, Zhang, Yu, Li, Ximing, Tao, Liang, & Bond, Z. S. (2012). Effects of speaker variability and noise on Mandarin fricative identification by native and non-native listeners. *The Journal of the Acoustical Society of America*, 132(2), 1130-1140. <https://doi.org/10.f385d3>
- Lin, Yen-Hwei. (2007). *The Sounds of Chinese*. Cambridge University Press.
- Liu, Zhao. (2019). *Análisis de las obstruyentes en chino y en español como L3. Estudio acústico y perceptivo para la categorización de errores* [Doctoral dissertation]. Universidad Autónoma de Barcelona.
- Lu, Yu-An. (2014). Mandarin fricatives redux: The psychological reality of phonological representations. *Journal of East Asian Linguistics*, 23(1), 43-69. <https://doi.org/10.1007/s10831-013-9111-5>
- Machuca, María Jesús, & Ríos, Antonio. (2017). *Frecuencia de fonemas y alófonos del español en la lengua oral espontánea* (Victoria Marrero & Eva Estabas, Eds.; pp. 93-97).
- Machuca, María Jesús. (2000). Articulación y pronunciación del español. En Santiago Alcoba (Ed.), *La expresión oral* (pp. 35-70). Ariel.
- Manrique, Ana María Borzone, & Massone, María Ignacia. (1981). Acoustic analysis and perception of Spanish fricative consonants. *The Journal of the Acoustical Society of America*, 69(4), 1145-1153. <https://doi.org/10.1121/1.385694>

- Martínez Celrá, Eugenio. (1984). *La fonética*. Teide.
- Martínez Celrá, Eugenio. (1996). *El sonido en la comunicación humana*. Octaedro.
- Mielke, Jeff. (2012). A phonetically based metric of sound similarity. *Lingua*, 122(2), 145-163. <https://doi.org/10.1016/j.lingua.2011.04.006>
- Ortega-García, Javier, Gonzalez-Rodríguez, Joaquín, & Marrero-Aguiar, Victoria. (2000). AHUMADA: A large speech corpus in Spanish for speaker characterization and identification. *Speech Communication*, 10. [https://doi.org/10.1016/s0167-6393\(99\)00081-3](https://doi.org/10.1016/s0167-6393(99)00081-3)
- Patil, Indrajeet. (2021). Visualizations with statistical details: The «ggstatsplot» approach. *Journal of Open Source Software*, 6(61), 3167. <https://doi.org/10.21105/joss.03167>
- Pulleyblank, Edwin G. (1984). *Middle Chinese: A study in historical phonology*. University of British Columbia Press.
- Quilis, Antonio, & Fernández, Joseph A. (1985). *Curso de fonética y fonología españolas*. Consejo Superior de Investigaciones Científicas.
- Real Academia Española – Asociación de Academias de la Lengua Española. (2011). *Nueva gramática de la lengua española. Fonética y fonología*. Espasa Calpe.
- RStudio Team. (2022). *Rstudio: Integrated Development Environment for R* [Manual]. <http://www.rstudio.com/>
- Stokes, Stephanie F., & Zhen, Fang. (1998). An electropalatographic description of Putonghua fricatives and affricates. *Asia Pacific Journal of Speech, Language and Hearing*, 3(2), 69-78. <https://doi.org/10.1179/136132898805577232>
- Sypiańska, Jolanta, & Constantin, Elena-Raluca. (2021). New vs. similar sound production accuracy: The uneven fight. *Yearbook of the Poznan Linguistic Meeting*, 7(1), 155-179. <https://doi.org/10.14746/yplm.2021.7.7>
- Trísková, Hana. (2008). Triskova, 2008 Are the Mandarin retroflex initial consonants really retroflex Are the palatals really palatal... The notes on terminology.pdf. *Oriental Institute of the Czech Academy of Sciences*.
- Zhang, Wei. (2010). *Non-native speakers speak in phonemes: A phono-acoustic analysis of fricatives and affricates by native and Chinese speakers of English* [Doctoral dissertation]. Purdue University.

Anexo

Tabla 4. Valores de significación de los análisis estadísticos

Fonema	Parámetro	Análisis	Resultado
/f/	Duración	Lengua china	t_{Yuen} (68.93) = 6.60, p = 6.93e-09 , Hedges's g = -1.38, CI _{95%} [-2.17, -0.71].
		Lengua española	t_{Yuen} (40.2) = 2.95, p = 5.24e-03 , Hedges's g = 0.73, CI _{95%} [0.35, 1.31].
		Grupo CH	t_{Yuen} (61.66) = 1.50, p = 0.14, Hedges's g = -0.33, CI _{95%} [-0.80, -0.03].
		Grupo ES	t_{Yuen} (37.27) = 6.81, p = 4.92e-08 , Hedges's g = 1.67, CI _{95%} [0.90, 2.93].
	Pico máximo	Lengua china	t_{Yuen} (34.23) = 5.53, p = 3.40e-06 , Hedges's g = -0.85, CI _{95%} [0.53, 1.60].
		Lengua española	t_{Yuen} (51.9) = 2.46, p = 0.02 , Hedges's g = 0.48, CI _{95%} [0.07, 0.80].
		Grupo CH	t_{Yuen} (61.99) = 0.75, p = 0.46, Hedges's g = 0.16, CI _{95%} [-0.19, 0.81].
		Grupo ES	t_{Yuen} (24.09) = 1.09, p = 0.29, Hedges's g = -0.54, CI _{95%} [-1.65, 0.28].
	Int. media	Lengua china	t_{Yuen} (66.62) = 6.73, p = 4.63e-09 , Hedges's g = 1.84, CI _{95%} [0.95, 3.08].
		Lengua española	t_{Yuen} (51.26) = 2.52, p = 0.01 , Hedges's g = 0.50, CI _{95%} [0.02, 1.28].
		Grupo CH	t_{Yuen} (57.68) = 0.13, p = 0.90, Hedges's g = 0.03, CI _{95%} [-0.40, 0.47].
		Grupo ES	t_{Yuen} (59.87) = 3.6, p = 6.52e-04 , Hedges's g = -0.61, CI _{95%} [-1.19, -0.24].
	SD	Lengua china	t_{Yuen} (56.16) = 6.30, p = 4.84e-08 , Hedges's g = 1.14, CI _{95%} [0.66, 1.76].
		Lengua española	t_{Yuen} (50.66) = 0.21, p = 0.84, Hedges's g = -0.04, CI _{95%} [-0.41, 0.34].
		Grupo CH	t_{Yuen} (50.78) = 1.85, p = 0.07, Hedges's g = 0.55, CI _{95%} [0.07, 1.17].
		Grupo ES	t_{Yuen} (37.49) = 3.35, p = 1.86e-03 , Hedges's g = -0.82, CI _{95%} [-1.52, -0.31].
	Asimetría	Lengua china	t_{Yuen} (46.76) = 8.08, p = 2.03e-10 , Hedges's g = -1.36, CI _{95%} [-2.14, -0.95].
		Lengua española	t_{Yuen} (52.09) = 0.56, p = 0.58, Hedges's g = -0.10, CI _{95%} [-0.55, 0.16].
		Grupo CH	t_{Yuen} (54.59) = 1.89, p = 0.06, Hedges's g = -0.51, CI _{95%} [-1.18, -0.06].
		Grupo ES	t_{Yuen} (31.63) = 3.89, p = 4.87e-04 , Hedges's g = -1.12, CI _{95%} [0.59, 1.89].

/f/	Curtosis	Lengua china	$t_{Yuen} (68.7) = 5.77, p = \mathbf{2.07e-07}$, Hedges's $g = -1.49$, $CI_{95\%} [-2.54, -0.71]$.
		Lengua española	$t_{Yuen} (39.86) = 1.12, p = 0.27$, Hedges's $g = 0.17$, $CI_{95\%} [-0.11, 0.38]$.
		Grupo CH	$t_{Yuen} (42.64) = 1.36, p = 0.18$, Hedges's $g = -0.53$, $CI_{95\%} [-1.23, 0.15]$.
		Grupo ES	$t_{Yuen} (60.83) = 5.61, p = \mathbf{5.27e-07}$, Hedges's $g = 0.85$, $CI_{95\%} [0.65, 1.25]$.
	COG	Lengua china	$t_{Yuen} (56.14) = 7.29, p = \mathbf{1.15e-09}$, Hedges's $g = 1.32$, $CI_{95\%} [0.81, 1.96]$.
		Lengua española	$t_{Yuen} (53) = 0.47, p = 0.64$, Hedges's $g = 0.09$, $CI_{95\%} [-0.27, 0.39]$.
		Grupo CH	$t_{Yuen} (55.92) = 1.68, p = 0.10$, Hedges's $g = 0.44$, $CI_{95\%} [0.01, 1.01]$.
		Grupo ES	$t_{Yuen} (35.53) = 3.65, p = \mathbf{8.44e-04}$, Hedges's $g = -0.93$, $CI_{95\%} [-1.52, -0.29]$.
/s/	Duración	Lengua china	$t_{Yuen} (40.75) = 4.25, p = \mathbf{1.22e-04}$, Hedges's $g = -0.91$, $CI_{95\%} [-1.52, -0.57]$.
		Lengua española	$t_{Yuen} (468.97) = 6.26, p = \mathbf{8.95e-10}$, Hedges's $g = 0.46$, $CI_{95\%} [0.33, 0.64]$.
		Grupo CH	$t_{Yuen} (27.87) = 3.86, p = \mathbf{6.16e-04}$, Hedges's $g = 0.71$, $CI_{95\%} [0.26, 1.44]$.
		Grupo ES	$t_{Yuen} (38.12) = 15.23, p = \mathbf{0.00}$, Hedges's $g = 2.73$, $CI_{95\%} [1.84, 3.97]$.
	Pico máximo	Lengua china	$t_{Yuen} (41.31) = 1.48, p = 0.15$, Hedges's $g = 0.32$, $CI_{95\%} [-0.08, 2.60]$.
		Lengua española	$t_{Yuen} (317.02) = 0.21, p = 0.83$, Hedges's $g = -0.01$, $CI_{95\%} [-0.11, 0.09]$.
		Grupo CH	$t_{Yuen} (23.95) = 1.09, p = 0.29$, Hedges's $g = 0.19$, $CI_{95\%} [-0.13, 2.07]$.
		Grupo ES	$t_{Yuen} (25.87) = 1.02, p = 0.32$, Hedges's $g = -0.17$, $CI_{95\%} [-0.49, 0.20]$.
	Int. media	Lengua china	$t_{Yuen} (28.11) = 1.15, p = 0.26$, Hedges's $g = 0.21$, $CI_{95\%} [-0.04, 0.76]$.
		Lengua española	$t_{Yuen} (468.2) = 7.63, p = \mathbf{1.34e-13}$, Hedges's $g = 0.55$, $CI_{95\%} [0.38, 0.65]$.
		Grupo CH	$t_{Yuen} (21.95) = 1.60, p = 0.12$, Hedges's $g = -0.28$, $CI_{95\%} [-0.60, 0.08]$.
		Grupo ES	$t_{Yuen} (30.78) = 3.00, p = \mathbf{5.26e-03}$, Hedges's $g = -0.51$, $CI_{95\%} [-0.99, -0.20]$.
	SD	Lengua china	$t_{Yuen} (31.6) = 1.59, p = 0.12$, Hedges's $g = 0.31$, $CI_{95\%} [-0.06, 0.70]$.
		Lengua española	$t_{Yuen} (442.74) = 1.15, p = 0.25$, Hedges's $g = 0.10$, $CI_{95\%} [-0.11, 0.24]$.
		Grupo CH	$t_{Yuen} (21.96) = 0.25, p = 0.81$, Hedges's $g = -0.04$, $CI_{95\%} [-0.45, 0.34]$.
		Grupo ES	$t_{Yuen} (31.86) = 3.25, p = \mathbf{2.74e-03}$, Hedges's $g = -0.56$, $CI_{95\%} [-0.86, -0.13]$.
	Asimetría	Lengua china	$t_{Yuen} (30.99) = 3.10, p = \mathbf{4.06e-03}$, Hedges's $g = -0.59$, $CI_{95\%} [-1.11, -0.25]$.
		Lengua española	$t_{Yuen} (383.01) = 1.52, p = 0.13$, Hedges's $g = -0.09$, $CI_{95\%} [-0.20, 0.02]$.
		Grupo CH	$t_{Yuen} (22.88) = 1.29, p = 0.21$, Hedges's $g = -0.23$, $CI_{95\%} [-0.54, 0.12]$.
		Grupo ES	$t_{Yuen} (27.92) = 3.51, p = \mathbf{1.55e-03}$, Hedges's $g = 0.58$, $CI_{95\%} [-0.20, 1.09]$.
	Curtosis	Lengua china	$t_{Yuen} (30.58) = 1.38, p = 0.18$, Hedges's $g = -0.26$, $CI_{95\%} [-0.20, 0.55]$.
		Lengua española	$t_{Yuen} (455.42) = 2.83, p = \mathbf{4.83e-03}$, Hedges's $g = 0.19$, $CI_{95\%} [0.07, 0.32]$.
		Grupo CH	$t_{Yuen} (22.88) = 1.56, p = 0.13$, Hedges's $g = 0.27$, $CI_{95\%} [-0.08, 0.61]$.
		Grupo ES	$t_{Yuen} (29.62) = 1.39, p = 0.18$, Hedges's $g = 0.23$, $CI_{95\%} [-0.13, 0.54]$.
	COG	Lengua china	$t_{Yuen} (34.54) = 2.36, p = \mathbf{0.02}$, Hedges's $g = 0.47$, $CI_{95\%} [0.14, 1.46]$.
		Lengua española	$t_{Yuen} (468.6) = 5.12, p = \mathbf{4.55e-07}$, Hedges's $g = 0.37$, $CI_{95\%} [0.19, 0.51]$.
		Grupo CH	$t_{Yuen} (23.03) = 1.82, p = 0.08$, Hedges's $g = 0.32$, $CI_{95\%} [-0.06, 0.87]$.
		Grupo ES	$t_{Yuen} (31.21) = 1.04, p = 0.31$, Hedges's $g = 0.18$, $CI_{95\%} [-0.19, 0.59]$.

Textos empleados para recoger muestras en este estudio

- Texto 1: El año nuevo

今天是春节，中国新年，也是一个传统佳节。高记者从欧洲回中国探望她的家人。爸爸在门上贴红色的对联，妈妈早准备了一桌饭菜。吃完以后，大家一起聊天看新年晚会。等到十二点整，全家人准备在院子里放鞭炮。因为这是一个重要习俗，去旧迎新。

- Texto 2: El otoño

冬天就要来了，浓浓的秋意中，一簇簇菊花层层落落绽放在清晨的校园中。一缕缕刺骨的秋风匆匆地走过，擦着老树的枯藤叶子，肆意地发出簌簌的声音。干枯的树木也被逼迫着散播冬日即将到来的消息。可突如其来得严寒让人措不及防。在这曾经漫天春意的园子里迷了路，更是足足增添了几分孤单伤感呢。或许那发怒了的秋老虎，也是个女子吧。品着一壶好茶，配着一支好歌，偌大的世间，我们只是一朵浮萍。

- Texto 3: Bruyninckx et al., 1994

El joyero Federico Vanero ha sido condenado por la audiencia de Santander a ocho meses de arresto mayor y cincuenta mil pesetas de multa por un delito de compra de objetos robados. La vista oral se celebró el miércoles pasado y, durante ella, uno de los fiscales, Carlos Valcárcel, pidió para el joyero tres años de prisión menor y una multa de cincuenta mil pesetas. Gracias a las revelaciones de Vanero de hace dos años y medio se llegó a descubrir la existencia de una sospechosa mafia policial en España, parte de la cual se vio envuelta en el llamado "caso el Nani".

- Texto 4: Ortega-García et al., 2000

Hay algo ahí, en el aire, que cambia el sentido de las cosas. Ese viento suave vuela, te toca la cara, mientras cuentas las hojas de los árboles. El agua corre buscando los campos. Al abrir las puertas de mi casa pienso: este país, una mañana más. A mi edad, comienzan a faltarme las fuerzas, ya casi no soy joven, y la muerte de mi mujer en la guerra me pesa mucho. Cuando el cuerpo llega a esa hora, la ciencia de los doctores no logra detener el paso del tiempo. De niño, allá en mi tierra, solía pasarme los días revolviendo de un lado a otro. Poco a poco, los coches de la ciudad fueron llamando mi atención. Mi madre decía que tuviera cuidado, pero yo me creía muy mayor, así que no tenía ni interés ni tiempo para mi propio signo. Pero sigo, es cierto, cuántas cosas buenas encontré entre su gente. Siuento los queridos veranos de entonces, no son siete, ni nueve, ni veinte. Debe ser que soy niño de nuevo en este cuerpo triste.