

Evaluación de varios alófonos de los fonemas españoles /f/, /x/, /l/ y /r/ por los hablantes nativos de español

Takuya KIMURA

1. INTRODUCCIÓN

Para los hablantes nativos de japonés que aprenden español, resulta difícil la distinción entre los fonemas /f/ y /x/ (especialmente cuando están seguidos de /u/ o /w/) y entre /l/ y /r/ (= [r]), tanto en la producción como en la percepción. Esta dificultad está causada por las diferencias del sistema consonántico de las dos lenguas.

Además del hecho de que en japonés no exista la consonante [f] ni la [x] como sonidos, este idioma solo tiene un fonema fricativo no sibilante, /h/, cuyos alófonos [h], [ç] y [ϕ] constituyen una distribución complementaria. Lo que merece notar aquí es que la secuencia japonesa /hu/ se realiza [ϕu] (p.ej. /hune/ [ϕune] “barco”), y que la secuencia /hw/, que solo se encuentra en palabras préstamos, se realiza [ϕ] (p.ej. /hwoHku/ [ϕo:kɯ] “tenedor” (< *ing.* fork)). De ahí que para los hablantes japoneses resulte difícilísima la distinción entre las secuencias españolas /fu/ (o /fw/) y /xu/ (o /xw/).

En cuanto a /l/ y /r/, es muy conocida la falta de distinción entre los sonidos [l] y [r] en japonés, ya que en este idioma hay un solo fonema líquido, normalmente escrito /r/, cuyos alófonos [l] y [r] muchas veces aparecen al azar en los mismos contextos. Por ejemplo, unos pronuncian la palabra /sara/ “plato” como [sala] y otros [sara], e incluso una misma persona puede pronunciar esta palabra de las dos maneras sin darse cuenta. Por lo tanto, se necesitan instrucciones precisas y entrenamiento adecuado para que los estudiantes japoneses adquieran la capacidad de distinguir estas dos consonantes.

Por otra parte, hay una teoría, cada vez más prevaleciente en el mundo de lingüística aplicada, que manifiesta que en la comunicación verbal no se requiere necesariamente la capacidad de pronunciar exactamente igual que los nativos hablantes de la lengua meta, sino la de pronunciar lo bastante bien como para no causar malentendidos ni dificultar la comunicación.¹⁾

En el presente trabajo informamos sobre los resultados de un experimento llevado a cabo por el autor cuyo objeto es saber qué consonantes son percibidas como alófonos normales de cada uno de los cuatro fonemas en cuestión.

2. EXPERIMENTO

El experimento se realizó en noviembre de 2014 en un aula normal del Centro Cultural Hispano-Japonés de la Universidad de Salamanca (España). Participaron 50 habitantes de Salamanca (40 mujeres y 10 hombres, de 18 a 76 años de edad, con un promedio de 29,7 años), cuya lengua materna es el español peninsular, y que no tenían experiencia con el japonés.

2.1. Materiales

Se elaboró una serie de ficheros de audio digital en formato WAV, cada uno de los cuales contenía la grabación por el autor de este artículo de una de las 12 secuencias de sonidos que se muestran en la Tabla 1, todas insertadas en la oración portadora *Digo _____ después.*²⁾ En lo que sigue, no se marcará el acento en las transcripciones fonéticas.

Tabla 1. Contenido de los estímulos utilizados.

(1)	Digo [fweɣo] después.	(6)	Digo [ola] después.
(2)	Digo [hweɣo] después.	(7)	Digo [o̞la] después.
(3)	Digo [ɸweɣo] después.	(8)	Digo [ora] después.
(4)	Digo [θweɣo] después.	(9)	Digo [o̞a] después.
(5)	Digo [xweɣo] después.	(10)	Digo [õ̞a] después.
		(11)	Digo [ora] después.
		(12)	Digo [o̞.ɾa] después.

Los signos fonéticos [l] y [l̃], de (7) y (10) respectivamente, indican consonantes acortadas digitalmente a base de la grabación de (6) [ola] y (9) [o̞a]. En la Figura 1 mostramos una parte de la forma de onda y el espectrograma del estímulo (6) y del (7) para su comparación. Se grabó primero la oración (6), en cuya grabación la duración de la consonante [l] fue de 83 ms, y después se quitó una parte de la misma consonante para hacerla más corta y obtener así el estímulo (7), en el que la duración de [l] es de 27 ms. Del mismo modo preparamos el estímulo (10) acortando la duración de la consonante retrofleja [l̃] (93 ms) del estímulo (9). La duración de [l̃] (estímulo 10) es de 27 ms.

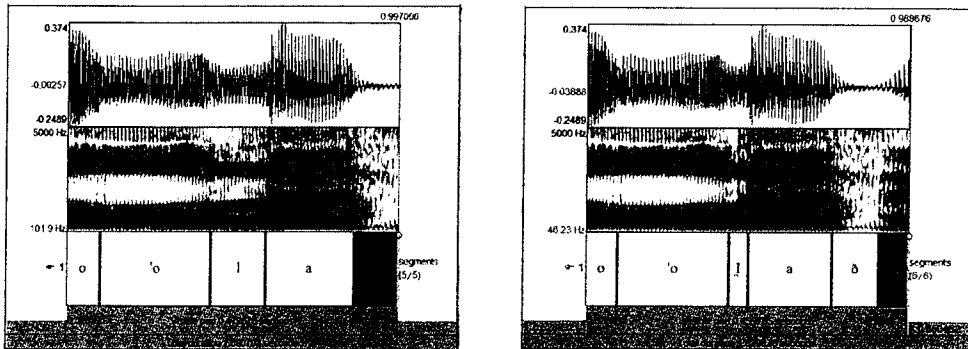


Figura 1. Oscilogramas y espectrogramas parciales del estímulo (6) con [l] (izquierda) y del (7) con [d] (derecha).

El motivo de este acortamiento es ver los posibles efectos de la duración consonántica al hacer que la duración de las aproximantes laterales fuera comparable a la de la vibrante simple [r]. Mostramos en la Figura 2 la parte que contiene la [r] del estímulo (8) para su comparación. La duración de la [r] en este estímulo es de tan solo 18 ms.

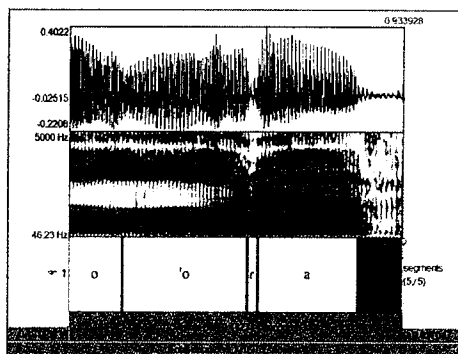


Figura 2. Oscilogramas y espectrogramas parciales del estímulo (8) con [r].

2.2. Procedimiento

Para la prueba de percepción, los 50 participantes escucharon los estímulos descritos a través de auriculares y evaluaron el nivel de la pronunciación de secuencias parecidas como la palabra *fuego* /fwego/ y la palabra *juego* /xwego/ en el caso de los estímulos (1) – (5), y como la palabra *hola* /ola/ y la palabra *hora* /ora/ en el caso de los estímulos (6) – (12).

Esta prueba forma parte de un experimento de mayores dimensiones con 79 estímulos en total,

sin embargo y para dar cuenta detallada de los resultados, el presente trabajo solamente informa sobre los 12 estímulos listados en la Tabla 1, ya que el resto de estímulos se presentarán en futuros artículos. El experimento constó de dos sesiones, en cada una de las cuales se presentaron las 79 oraciones en un orden aleatorio. Estos órdenes también fueron diferentes en la primera sesión y en la segunda. Cada estímulo se presentó dos veces seguidas. A los sujetos, nativos de español, se les pidió que evaluaran la pronunciación de la palabra objetivo utilizando una escala de 1 a 5, siendo 1 «la pronunciación es tan mala que resulta ininteligible o suena como otra palabra distinta a la escrita» y 5 «la pronunciación es tan buena como la de un hablante nativo de español». En los análisis que siguen, la puntuación total de las dos evaluaciones para un mismo estímulo será considerado como la evaluación de ese estímulo, por lo que la valoración máxima de cada juez será de un 10, y el mínimo de un 2.

La Figura 3 muestra el formato de la hoja de respuestas que se utilizó en el experimento. Cada uno de los estímulos apareció dos veces bajo diferentes palabras escritas. Por ejemplo, el mismo estímulo [ora] se presentó en «2. hora» y en «8. hola», por lo que los sujetos tuvieron que juzgar el nivel de pronunciación de este estímulo como la palabra meta *hora* en el número 2, y como la palabra meta *hola* en el número 8.

	(peor)				(mejor)
2. hora	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
	(. . .)				
4. juego	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
	(. . .)				
6. fuego	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
	(. . .)				
8. hola	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
	(. . .)				
49. hola	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
	(. . .)				

Figura 3. Formato de la hoja de respuestas.

2.3. Resultados

2.3.1. *fuego / juego*

Para calibrar objetivamente la importancia de las diferentes valoraciones obtenidas se aplicó el análisis de varianza de un factor (*one-way ANOVA*), con el tipo de consonante como el factor, a los puntos de evaluación

dados por los jueces a cada una de las consonantes [f], [h], [ϕ], [θ] y [x] grabadas como posibles alófonos de la /f/ de *fuego* (Tabla 1). El resultado mostró diferencias significativas en los distintos tipos de consonantes ($F_{4, 196} = 1102,81, p < 0,001$).³⁾

En la columna central de la Tabla 2 se muestran en orden descendente los puntos de evaluación de cada consonante como alófonos de /f/ en *fuego*, mientras que en la columna derecha aparecen los resultados de comparación múltiple aplicada a cada pareja de consonantes colindantes. (El ajuste de los valores *p* está basado en Benjamini & Hochberg, 1995).

Las cifras de puntos de evaluación que aparecen en la Tabla 2 y en todas las tablas que siguen son valores de promedio de los 50 sujetos. Las cifras de la columna derecha son los valores *p*, resultados de la prueba *t* para comprobar si la diferencia de los promedios de consonantes colindantes es estadísticamente significativa en cada caso. Los valores *p* se acompañan de signos que dejan ver fácilmente la significación: *** ... $p < 0,001$, ** ... $p < 0,01$, * ... $p < 0,05$ (todos estos indican que la diferencia es estadísticamente significativa); + ... $p < 0,10$ (en este caso se dice que la diferencia es marginalmente significativa); n.s. ... $p \geq 0,10$ (esto quiere decir que no hay diferencia estadísticamente significativa).

Tabla 2. Puntos de evaluación de cada consonante como alófono del /f/

[f]	9,66	$p = 0,018$ (*)
[θ]	9,20	
[ϕ]	9,06	$p = 0,485$ (n.s.)
[h]	2,18	$p < 0,001$ (***)
[x]	2,12	$p = 0,583$ (n.s.)

Salta a la vista la alta evaluación de la consonante [θ], junto con la [ϕ], aunque hay una diferencia significativa entre la valoración de estas dos consonantes y la [f], que es la mejor valorada.⁴⁾

A continuación, se aplicó el análisis de varianza de un factor con el tipo de consonante como el factor, a los puntos de evaluación dados a cada consonante como alófonos de /x/ en la palabra *juego*. Nuevamente el resultado mostró diferencias significativas en los distintos tipos de consonantes ($F_{4, 196} = 363,19, p < 0,001$).

En la columna central de la Tabla 3 se muestran en orden descendente los puntos de evaluación de cada consonante como alófonos de /x/ en *juego*.

Tabla 3. Puntos de evaluación de cada consonante como alófono del /x/

[x]	9,36	$p < 0,001$ (***)
[h]	5,78	
[f]	2,36	$p < 0,001$ (***)
[ϕ]	2,26	
[θ]	2,26	$p = 0,551$ (n.s.)
		$p = 1,000$ (n.s.)

La consonante [x] es el único alófono que obtuvo una alta evaluación. También se debe destacar que la consonante aspirada [h] obtuvo una evaluación mediana (5,78), aunque la diferencia con la [x] es grande.⁵⁾ Sin embargo, las consonantes [f], [ϕ] y de [θ] presentan puntuaciones muy bajas.

La Tabla 4 muestra el resultado de la prueba *t*, aplicada a cada una de las cinco consonantes comparando los puntos de evaluación otorgados cuando la meta era /f/ (*fuego*) y cuando era /x/ (*juego*).

Tabla 4. Comparación de los puntos de evaluación como /f/ y como /x/

	/f/	<i>t</i> (df = 49)	<i>p</i>	/x/
[f]	9,66	37,98	< 0,001 (***)	2,36
[θ]	9,20	25,69	< 0,001 (***)	2,26
[ϕ]	9,06	33,33	< 0,001 (***)	2,26
[h]	2,18	13,59	< 0,001 (***)	5,78
[x]	2,12	43,74	< 0,001 (***)	9,36

Se ve que las cinco consonantes muestran una diferencia significativa muy grande en los puntos obtenidos dependiendo de si la meta para los jueces era /f/ o /x/. Como hemos mencionado, la consonante [x] es la única que obtiene una alta evaluación como un alófono del fonema /x/, pero la evaluación de la [h] también es mucho más alta como /x/ que como /f/.

2.3.2. *hola / hora*

En primer término se aplicó el análisis de varianza de un factor, con el tipo de consonante como el factor, a los puntos de evaluación dados a cada una de las consonantes [l], [ɫ], [ʌ], [ʌ̃], [ɹ], [r] y [r̃] como alófonos del /l/ en la palabra *hola*. El resultado mostró que existen diferencias significativas en los distintos tipos de consonantes ($F_{6, 294} = 354,82, p < 0,001$).

En la columna central de la Tabla 5 se muestran en orden descendente los puntos de evaluación

de cada consonante como alófonos del /l/ de *hola*. En la columna derecha aparecen los resultados de la comparación estadística múltiple aplicada a cada pareja de consonantes colindantes en esa gradación.

Tabla 5. Puntos de evaluación de cada consonante como alófono del /l/

[l]	9,62	$p = 0,003 (**)$
[l̥]	9,12	
[ɫ]	8,46	$p < 0,001 (***)$
[l̃]	8,02	$p = 0,131 (n.s.)$
[ɫ̃]	3,64	$p < 0,001 (***)$
[r]	2,20	$p < 0,001 (***)$
[r̃]	2,10	$p = 0,302 (n.s.)$

Un hecho inesperado de difícil explicación observado en este resultado es que la evaluación de [l̥] es ligera aunque significativamente más alta que la de [l]. Es posible que sea un problema específico de los estímulos utilizados. Aún así, la consonante esperada [l] obtiene un valor muy alto. Las dos consonantes acortadas obtienen evaluaciones bastante altas, pero ambas son significativamente más bajas que las de las no acortadas.

A continuación, se aplicó el análisis de varianza de un factor, con el tipo de consonante como el factor, a los puntos de evaluación dados a cada consonante como alófonos del /r/ de la palabra *hora*. Otra vez el resultado mostró diferencias significativas en los distintos tipos de consonantes ($F_{6, 294} = 184,35, p < 0,001$).

En la columna central de la Tabla 6 se muestran en orden descendente los puntos de evaluación de cada consonante como alófonos del /r/ de *hora*.

Tabla 6. Puntos de evaluación de cada consonante como alófono del /r/

[r]	8,90	$p < 0,001 (***)$
[r̃]	3,90	
[ɹ]	3,48	$p = 0,247 (n.s.)$
[r̃]	2,68	$p < 0,001 (***)$
[ɹ̃]	2,62	$p = 0,762 (n.s.)$
[l]	2,28	$p = 0,068 (+)$
[l̃]	2,18	$p = 0,523 (n.s.)$

Aunque no se observa una diferencia significativa en ninguna de las parejas de dos consonantes colindantes desde la [r] del cuarto puesto hasta la [ʎ] del séptimo, sí existe una diferencia significativa entre la cuarta [r] y la sexta [ʎ] ($p = 0,010$ (*)), y entre la quinta [ʎ] y la séptima [ʎ] ($p = 0,005$ (**)).

Hay que mencionar que la aproximante lateral acortada [l̥] obtiene una evaluación comparable con la de [ɹ]. También salta a la vista la gran diferencia entre la evaluación de [l̥] y la de [ʎ]. Este resultado sugiere la posibilidad de que las consonantes líquidas, cuando son lo suficientemente cortas, se identifiquen como alófonos del fonema /r/ independientemente de si son laterales o no. Todas las otras aproximantes laterales obtienen una evaluación baja. La [r] es el alófono normativo de un fonema diferente, que es el /r/, lo que hace comprensible la baja evaluación que dicha consonante ha obtenido.

La Tabla 7 muestra el resultado de la prueba *t* aplicada a cada una de las cinco consonantes que compara la puntuación obtenida cuando se esperaba /l/ (*hola*) y cuando se esperaba /r/ (*hora*).

Tabla 7. Comparación de los puntos de evaluación como /l/ y como /r/

	/l/	<i>t</i> (df = 49)	<i>p</i>	/r/
[ʎ]	9,62	58,00	< 0,001 (***)	2,18
[ʎ]	9,12	38,65	< 0,001 (***)	2,28
[ʎ]	8,46	24,72	< 0,001 (***)	2,62
[l̥]	8,02	8,14	< 0,001 (***)	3,90
[ɹ]	3,64	0,44	0,659 (n.s.)	3,48
[r]	2,20	32,44	< 0,001 (***)	8,90
[r]	2,10	4,05	< 0,001 (***)	2,68

Se observa una clara tendencia a que las aproximantes laterales se identifiquen como la consonante /l/ y a que la vibrante simple [r] se identifique como /r/. El sonido [ɹ], normalmente clasificado como un sonido rótico en fonética general, no parece que sea una realización normal de ninguno de los dos fonemas para los hispanohablantes.

3. A MODO DE CONCLUSIÓN

A través de este experimento se han podido obtener varios resultados de interés:

1. Se demuestra que los dos sonidos [θ] y [ϕ] son aceptados casi igualmente como alófonos del fonema /f/.
2. La aproximante lateral retrofleja acortada puede ser identificada como alófono del fonema /r/.

concretamente con la variante simple de la palabra *hora*.

3. La consonante aspirada [h] se ha valorado positivamente como sustituta de la fricativa velar, ya que efectivamente este fonema se produce aspirado en muchas variedades del español.
4. Como era de esperar, la vibrante múltiple ha obtenido una baja puntuación para sustituir a la simple en la palabra *hora*, dado que en ese contexto estas dos consonantes sí se oponen claramente en español (*caro / carro*), algo que incluso se refleja en la ortografía.

Por otra parte, ha podido haber algunos problemas técnicos en este experimento, ya que es imposible garantizar que las partes no correspondientes a las consonantes estudiadas, es decir, el resto de la palabra, fueran completamente iguales en todos los estímulos utilizados. Tampoco se puede descartar la posibilidad de que los estímulos preparados por el autor del presente trabajo, un fonetista japonés hablante de español como segunda lengua, contengan algunas características personales en la pronunciación de las consonantes objetivo.

Se espera un experimento más estrictamente controlado para elevar la fiabilidad de las informaciones aquí obtenidas.

Agradecimientos

Este trabajo es una parte de los resultados del proyecto titulado «Estudio sobre percepción del fonema español /y/, producción de las consonantes japonesas /y/ y /zy/, y sobre evaluación de la pronunciación española de estudiantes japoneses por hablantes nativos de español», aprobado por el Comité de Ética Referente a Investigaciones con Sujetos Humanos de la Universidad Sofía.

Este trabajo ha sido subvencionado por el subsidio especial para estudios en el extranjero del año escolar 2014/2015 de la Universidad Seisen.

Agradezco profundamente al profesor Takayuki Arai (Universidad Sofía, Tokio) su asesoramiento y su amable ayuda en varios aspectos, a todo el personal del Centro Cultural Hispano-Japonés de la Universidad de Salamanca su amable colaboración no solamente durante el experimento sino también en el previo reclutamiento de participantes, y, por supuesto, a todos los participantes en el experimento su colaboración y paciencia.

Mi agradecimiento va también a José Antonio Martín Gómez (Universidad de La Laguna), quien tuvo la amabilidad de examinar la versión provisional del manuscrito y aportó útiles comentarios y sugerencias para mejorarlo.

NOTAS

1) Por ejemplo según Gil (2007:97), “se ha dicho alguna vez que sólo los espías necesitan hablar una lengua extranjera sin acento alguno que los delate, mientras que a los restantes seres humanos, con formas de vida más vulgares, les basta casi siempre con resultar relativamente inteligibles en sus respectivas necesidades comunicativas.”

2) La grabación se llevó a cabo en julio de 2014, en la cabina insonorizada del Laboratorio del profesor Takayuki Arai, Facultad de Ciencia e Ingeniería, Universidad Sofía. Para la grabación utilizamos el grabador PCM linear LS-10 de Olympus y el micrófono condensador electret ECM-23F5 de Sony. La frecuencia de muestreo fue de 44,1 kHz y la resolución, 16 bit. Para las manipulaciones de sonido y la elaboración de gráficos utilizamos el software Praat, ver.5.3.82 (Boersma & Weenick, 2001).

3) Para las análisis estadísticas utilizamos los softwares R, ver.3.1.2. (R Core Team, 2014) y js-STAR 2012, release 2.0.7j (Tanaka & Nakano, 2013).

4) La alta evaluación de [θ] como alófono de /f/ se debe probablemente a una propiedad acústica que comparten la fricativa labiodental y la dental: una baja intensidad. Véanse Quilis (1981:229-231) y Martínez y Fernández (2013:105-108).

5) La evaluación relativamente alta de [h] como alófono de /x/ podría atribuirse a que esta consonante se usa en lugar de [x] en muchas variedades del español. La pronunciación [hweyo], aunque es poco común en Salamanca, debe de haber sonado bastante familiar.

REFERENCIAS BIBLIOGRÁFICAS

- Benjamini, Yoav & Hochberg, Yosef (1995), “Controlling the false discovery rate: a practical and powerful approach to multiple testing”, *Journal of the Royal Statistical Society, Series B* 57 (1): 289–300.
- Boersma, Paul & Weenink, David (2001), “Praat, a system for doing phonetics by computer”, *Glott International*, 5, 9-10, 341-345.
- Gil Fernández, Juana (2007), *Fonética para profesores de español: de la teoría a la práctica*, Arco Libros.
- Martínez Celdrán, Eugenio & Fernández Planas, Ana María (2013), *Manual de fonética española: Articulación y sonidos del español, 2.ª ed.*, Ariel.
- Quilis, Antonio (1981), *Fonética acústica de la lengua española*, Gredos.
- R Core Team (2014), *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>
- Tanaka, Satoshi & Nakano, Hiroyuki (2013), *Introducción a los análisis de datos con R & STAR* (en japonés), Shin-yosha.