

BRYGIDA SAWICKA-STEPÍŃSKA

ORCID: 0000-0002-2019-8275

Uniwersytet im. Adama Mickiewicza w Poznaniu

Correo: brygida.sawicka-stepinska@amu.edu.pl

Hacia una clasificación automatizada de los alófonos de /s/ en el español guayaquileño*

Palabras clave: debilitamiento de /s/ — aprendizaje automático — español guayaquileño — procesamiento digital de voz — fonética acústica.

Resumen

El presente artículo tiene como objetivo generar una manera de clasificar automáticamente los alófonos de /s/ retenidos y debilitados tras encontrar una correspondencia entre sus parámetros acústicos y el tipo cualitativo de la realización. El conjunto de datos utilizado para este estudio proviene del CHARG (Corpus de Habla Radiofónica de Guayaquil). Los parámetros considerados para la descripción acústica de la consonante fricativa en cuestión son: la duración (DUR), el centro de gravedad (COG) y el porcentaje de ensordecimiento (UNVOI). Esta elección se fundamenta en la ratificación de su valor indicativo de la lenición del fonema. Para llevar a cabo la tarea, se extrajeron los valores de los parámetros acústicos. Como siguiente paso, se seleccionaron aleatoriamente el 30% de las observaciones y se clasificaron perceptivamente como retenidas o debilitadas. El análisis estadístico reveló una correlación débil pero significativa entre los parámetros, y una correlación satisfactoria entre los valores de DUR/COG/UNVOI y los resultados de la evaluación perceptiva de las fricativas. El presente estudio es una etapa inicial para elaborar un algoritmo entrenado para la clasificación de consonantes en categorías discretas basadas en medidas acústicas objetivas, trabajando con grandes bases de datos obtenidos fuera de laboratorio.

1. Introducción

El debilitamiento de /s/ es un fenómeno ampliamente estudiado en la lingüística hispánica. Rafael Lapesa lo considera el cambio fonético más potente

* El presente estudio forma parte del proyecto NCN 2017/25/N/HS2/00006, financiado por el Centro Nacional de Ciencia de Polonia.

en la sincronía y la diacronía del español¹. Sería una tarea imposible llevar a cabo un estudio que proporcione cifras exactas sobre los hablantes que aspiran o eliden la /s/; sin embargo, se estima que el alcance de la lenición de /s/ cubre más de la mitad de las zonas de habla hispana².

En cuanto a la articulación, el debilitamiento de la /s/ (o aspiración, como se lo denomina en la tradición lingüística hispánica) es un ejemplo de debucalización, que significa que el lugar de articulación de la consonante se desplaza desde los alvéolos hacia la laringe³. El proceso se da para minimizar el esfuerzo articulatorio y en búsqueda de la sílaba abierta⁴. A grandes rasgos, la debucalización de /s/ resulta en un sonido aspirado [h] o una elisión total, especialmente en posición de coda, dependiendo, entre otros condicionantes, del contexto fónico, de factores diafásicos, diastráticos y de la variedad dialectal⁵.

En el presente artículo dejamos de lado los factores mencionados. Este estudio se concentra en las variables dependientes, en búsqueda de una descripción acústica satisfactoria que nos permita crear una herramienta para la clasificación automática de las realizaciones de /s/ en datos obtenidos fuera de laboratorio. Empleando un nuevo corpus de habla radiofónica de Guayaquil, hemos examinado la utilidad de tres parámetros acústicos: centro de gravedad (COG), duración (DUR) y porcentaje de ensordecimiento (UNVOI). Los parámetros nos sirven para medir el nivel de debilitamiento de /s/ en posición de coda, así como para clasificar las realizaciones del fonema en dos categorías discretas: debilitada frente a retenida. Estas tres variables han sido empleadas en estudios anteriores para describir y analizar consonantes fricativas⁶. El descenso de DUR y COG se considera un indicador de lenición, así como la sonorización, puesto que el movimiento de las cuerdas vocales debilita la corriente de aire indispensable para la producción de sonidos sibilantes⁷.

Los análisis acústicos instrumentales del debilitamiento de /s/ existen, pero son escasos en comparación con los que se basan en observaciones impresio-

¹ Cfr. M. de la Sierra Tejada Giráldez, *Convergencia y divergencia entre comunidades de habla: a propósito de la /s/ implosiva. Contribución al estudio de los patrones sociolingüísticos del español de Granada*, tesis doctoral, Granada, Universidad de Granada, 2015, p. 6.

² R.M. Hammond, 2001, cf. A. Walker, Ch. García, Y. Cortés y K. Campbell-Kibler, "Comparing social meanings across listener and speaker groups: The indexical field of Spanish /s/", *Language, Variation and Change*, 26, 2014, p. 172.

³ Real Academia Española, *Nueva gramática de la lengua española*, 2011, p. 198.

⁴ D.L. Canfield, *El español de América: Fonética*, Barcelona, Editorial Crítica, 1988, p. 17; V. Marrero, "Estudio acústico de la aspiración en español", *Revista de Filología Española*, LXX, 1990, p. 354.

⁵ K. Broś, *Survival of the Fittest: Fricative Lenition in English and Spanish from the Perspective of Optimality Theory*, Cambridge, Cambridge Scholars Publishing, 2015, p. 3.

⁶ E.K. Brown, "Paradigmatic Peer-Pressure: Word-Medial, Syllable-Initial /s/ Lenition in Dominican Spanish", en: S.M. Alvord (ed.), *Selected Proceedings of the 5th Conference on Laboratory Approaches to Romance Phonology*, Somerville, MA, Cascadilla Proceedings Project, 2011, pp. 46–58, o R.J. File-Muriel, E.K. Brown, "The gradient nature of s-lenition in Caleño Spanish", *Language, Variation and Change*, 23, 2011, pp. 223–243.

⁷ R.J. File-Muriel, E.K. Brown, *op. cit.*, pp. 228–229.

nistas. A modo de ejemplo, Marrero⁸ ofrece una detallada descripción acústica y articulatoria de tres tipos de sonidos aspirados (laríngeo, faríngeo y velar), poniendo en duda el sentido de analizar los segmentos por separado, puesto que su estructura armónica puede depender de las vocales vecinas. Aun así, las consonantes fricativas poseen un valor DAC (*Different Articulatory Constraints*) relativamente alto, lo cual quiere decir que son más resistentes a los efectos coarticulatorios que, p. ej., las oclusivas⁹. File-Muriel y Brown¹⁰ llevan a cabo un análisis de correlación entre los parámetros acústicos de /s/ y una serie de factores lingüísticos. Ambos estudios se realizaron sobre grabaciones de laboratorio, en relativamente pocos hablantes (cuatro en el estudio de Marrero y ocho en el estudio de File-Muriel y Brown).

Asimismo, queda demostrado que los tres parámetros en cuestión son portadores de rasgos distintivos que posibilitan la percepción y categorización de los alófonos de /s/. Cuanto más largo el sonido de frecuencias altas, más elevada es la probabilidad de que se clasifique como sibilante¹¹. Respecto al porcentaje de ensordecimiento, en la jerarquía de sonoridad, las consonantes sonorizadas se consideran más cercanas a las vocales y, por lo tanto, más débiles que los sonidos sordos en cuanto al esfuerzo articulatorio requerido para su producción¹². Nuestra hipótesis es que existe una relación entre estos tres parámetros acústicos y las categorías discretas del debilitamiento de /s/, la cual en el futuro nos servirá de base para clasificar los alófonos de /s/ de forma automática, evitando el riesgo de caer en errores producidos por la percepción humana.

El presente artículo da cuenta del estudio inicial que forma parte del proyecto para el análisis lingüístico y extralingüístico del debilitamiento de /s/ en el habla de locutores de radio en Guayaquil (Ecuador). Los datos utilizados vienen del CHARG (Corpus de Habla Radiofónica de Guayaquil). El español guayaquileño no goza de mucha atención por parte de los lingüistas. Aunque existen algunos estudios valiosos sobre el español ecuatoriano¹³, pocos de ellos tratan la variedad de la región del Guayas (vale la pena mencionar aquí un estudio de Estrella sobre el léxico de 2001), y menos todavía su fonética. De hecho, la referencia principal, *El español del Ecuador* de Humberto Toscano Mateus, data del año 1953. No sería sorprendente si no fuera por el hecho de que Guayaquil, la ciudad principal de la costa ecuatoriana, es actualmente el mayor centro urbano del país, con una población ligeramente mayor a la de

⁸ V. Marrero, *op. cit.*, p. 351.

⁹ Cfr. E. Martínez Celdrán, A.M. Fernández Planas, *Manual de fonética española. Articulaciones y sonidos del español*, 2^{da} edición, Barcelona, Ariel Letras. 2013, p. 102.

¹⁰ R.J. File-Muriel, E.K. Brown, *op. cit.*

¹¹ *Ibidem*, p. 230.

¹² G.N. Clements, "The role of the sonority cycle in core syllabification", en: J. Kingston, M.E. Beckman (eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, Cambridge, Cambridge University Press, 1990, pp. 283–333.

¹³ Cfr. los trabajos de Estrella, Flores Mejía, Palacios, Haboud o Strycharczuk *et al.*

la capital, Quito¹⁴. Las dos ciudades forman parte de dos zonas dialectales diferentes y la distancia entre ellas, tanto lingüística como de otra índole, está fuertemente arraigada en la conciencia social¹⁵. La creación del CHARG está motivada por la falta de descripciones satisfactorias de la variante del español hablada en la costa ecuatoriana.

La hipótesis principal del proyecto es que en la variedad guayaquileña del español, que se clasifica entre las variedades intermedias, las realizaciones debilitadas de /s/ van ganando un prestigio comparable al de las realizaciones sibilantes. Nuestro estudio tiene como objetivo elaborar un método para el análisis instrumental semi-automático de conjuntos de datos grandes que no hayan sido recogidos en condiciones de laboratorio.

2. Materiales y métodos

CHARG (Corpus de Habla Radiofónica de Guayaquil) es la primera base de datos de este tipo para la variedad del español hablada en Guayaquil. La creación del CHARG está motivada por el interés en la exploración lingüística y extralingüística del debilitamiento de /s/. Los estudios anteriores sobre otras variantes dialectales del español afirman que la radio puede ser una fuente de datos útil para este objetivo¹⁶.

La población meta del CHARG se define como el universo radiofónico de Guayaquil¹⁷. Se eligieron ocho emisoras locales populares. El tiempo de transmisión semanal se calculó para cada emisora. De ese cálculo, se excluyeron los fines de semana, puesto que su programación varía entre una semana y otra y contiene mucho material que no pudo entrar en el corpus debido a la baja calidad del sonido (p.ej. transmisiones en vivo, música programada, etc.). El tiempo de transmisión fue calculado proporcionalmente para obtener el total de 24 horas de grabaciones, lo cual da un promedio de tres horas por emisora. La ley ecuatoriana¹⁸ obliga a todas las emisoras a clasificar los programas con al menos una de estas seis etiquetas: informativo, de opinión, entretenimiento, formativo/educativo/cultural, deportivo y publicidad. Ya que

¹⁴ Instituto Nacional de Estadísticas y Censos, “Fascículo provincial Guayas”, 2010, [<http://www.ecuadorencifras.gob.ec/wp-content/descargas/Manu-lateral/Resultados-provinciales/guayas.pdf>]; “Fascículo provincial Pichincha”, 2010 <<http://www.ecuadorencifras.gob.ec/wp-content/descargas/Manu-lateral/Resultados-provinciales/guayas.pdf>>, 18 de septiembre de 2019.

¹⁵ E. Flores Mejía, “Actitudes lingüísticas en Ecuador: Una tradición normativa que subsiste”, en: A.B. Chiquito, M.Á. Quesada Pacheco (eds.), *Actitudes lingüísticas de los hispanohablantes hacia el idioma español y sus variantes*, Bergen, BeLLS, 2014, pp. 409–488.

¹⁶ T.L. Flores, “Velar palatalization in Chilean public speech”, *Glossa: a journal of general linguistics*, 1 (6), 2016, pp. 1–17.

¹⁷ A.M. López González, *El lenguaje radiofónico en la ciudad de Almería. Análisis Sociolingüístico*, tesis doctoral, Almería, Universidad de Almería, 2001, p. 81.

¹⁸ LOC (Ley Orgánica de Comunicación), Registro Oficial Suplemento 22, 25 de junio de 2013, art. 106.

la variable principal considerada para el estudio final es el tipo de contenido del programa, los tiempos de grabación fueron distribuidos proporcionalmente siguiendo este criterio.

Todas las grabaciones fueron recopiladas en agosto y septiembre de 2018 partiendo de las transmisiones en línea de acceso público y de algunos podcasts adicionales. Las transmisiones en línea se grabaron en el formato wav mediante el programa Foobar2000 con una frecuencia de muestreo de 44.1 Hz.

Los locutores incluidos en el corpus tienen al menos 18 años, son residentes nativos de Guayaquil y tienen educación superior. Como la mayoría de ellos son figuras públicas, la información biográfica usada para etiquetar a los hablantes se obtuvo de las páginas web de las emisoras, de los perfiles profesionales y personales de las redes sociales, de los artículos en línea y del contenido mismo de los programas. Todas las expresiones pronunciadas por los hablantes han sido transcritas ortográficamente (SAMPA) y fonéticamente mediante los programas Annotation Pro¹⁹, Praat²⁰ e EasyAlign²¹, siempre que la calidad de sonido fuera satisfactoria (los programas con música y ruido de fondo, es decir, un 13% del total del corpus, fueron sustituidos por otros que cumplieran con los mismos requisitos).

El análisis estadístico ha sido realizado en R (versión 3.5). Las distribuciones de COG, UNVOI y DUR se han visualizado con el estimador de densidad *kernel*, tal como se aplica en la función *stat_density* del paquete *ggplot2* de R. Para ajustar el clasificador, hemos usado regresión logística binaria estándar del paquete *nnet* de R.

3. Resultados

3.1. Conjunto de datos para el presente estudio

En este estudio se ha aplicado la mitad de la muestra de programas etiquetados bajo el estrato “informativo”. El conjunto de datos está compuesto de 3 horas y 11 minutos de grabaciones (el 13% del total del corpus) y contiene material hablado de 30 locutores (Tabla 1).

¹⁹ K. Klessa, M. Karpiński, A. Wagner, “Annotation Pro – a new software tool for annotation of linguistic and paralinguistic features”, en: D. Hirst, B. Bigi (eds.), *Proceedings of the Tools and Resources for the Analysis of Speech Prosody (TRASP) Workshop*, Aix-en-Provence, 2013, pp. 51–54.

²⁰ P. Boersma, D. Weenink, “Praat: doing phonetics by computer” [software]. Versión 6.0.29, 2017, <<http://www.praat.org/>>, 24 de mayo de 2017.

²¹ J.-Ph. Goldman, “EasyAlign: an automatic phonetic alignment tool under Praat”, en: *Proceedings of InterSpeech*, Firenze, 2011.

Tabla 1: Distribución de los hablantes según edad y sexo

	edad		
	50 < años	30–50 años	> 30 años
hombres	7	9	1
mujeres	2	9	2

Las grabaciones provienen de 19 programas de 7 emisoras.

Hemos transcrito ortográficamente (SAMPA) todas las expresiones emitidas por los presentadores de radio con el uso del programa Annotation Pro²². A continuación, los archivos fueron alineados automáticamente mediante EasyAlign²³ en tres pasos: macro-segmentación a nivel de expresión, conversión grafema-a-fonema y segmentación fónica. Hemos verificado y ajustado a mano las fronteras de todos los segmentos que contenían /s/ con el uso de Praat²⁴. La segmentación exigía una observación simultánea del espectrograma de banda ancha y del oscilograma. Para definir el *onset* y el *offset* de las consonantes, hemos seguido el procedimiento expuesto por Jongman *et al.*²⁵: para la /s/ sorda, hemos situado el *onset* en el punto donde por primera vez aparecían las altas frecuencias de energía y se observaba un crecimiento rápido de los cruces por cero, y el *offset* donde la frecuencia cambiaba en el espectrograma y el número de cruces por cero disminuía; para la /s/ sonorizada, hemos puesto el *onset* en el punto del primer cambio en el período de la onda y el *offset* donde se daba el número de cruces por cero característico para el período anterior.

3.2. Mediciones acústicas

Del total de las 2977 observaciones de /s/ en posición de coda, hemos excluido las secuencias de dos /s/, ya que es prácticamente imposible definir la frontera entre estos dos sonidos. Para medir los parámetros acústicos de /s/, hemos seguido el procedimiento propuesto por Brown²⁶. Mediante el código de Praat²⁷, hemos extraído la duración (DUR) en segundos, el centro de gravedad (COG) en hercios y el porcentaje de ensordecimiento (UNVOI) de las consonantes en cuestión. Para medir el COG, solo se ha aplicado el 60% en el medio del segmento para evitar la influencia de los segmentos colindantes.

²² K. Klessa *et al.*, *op. cit.*

²³ J.-Ph. Goldman, *op. cit.*

²⁴ P. Boersma, D. Weenink, *op. cit.*

²⁵ A. Jongman, R. Wayland, S. Wong, “Acoustic characteristics of English fricatives”, *Journal of the Acoustical Society of America*, 108, 2000, p. 1255.

²⁶ E.K. Brown, *op. cit.*, p. 51.

²⁷ Una modificación del código de E.K. Brown, 2014, <www-personal.ksu.edu/~ekbrown>, 18 de septiembre de 2019. El código está alojado en Zenodo <<https://doi.org/10.5281/zenodo.3972028>>.

3.3. Clasificación perceptiva

Para ejecutar la evaluación de las consonantes en cuestión, hemos extraído aleatoriamente un 30% de observaciones ($n = 898$) de cada hablante y las hemos clasificado perceptivamente como retenidas (sibilantes) o debilitadas (aspiradas). Nos hemos servido tanto de la percepción auditiva como del sonido visible para decidir entre las dos categorías. Hemos podido escuchar cada segmento un número de veces ilimitado para asegurar una evaluación lo más acertada posible. La Figura 1 proporciona la ilustración de un segmento clasificado como retenido y la Figura 2 de un segmento debilitado.

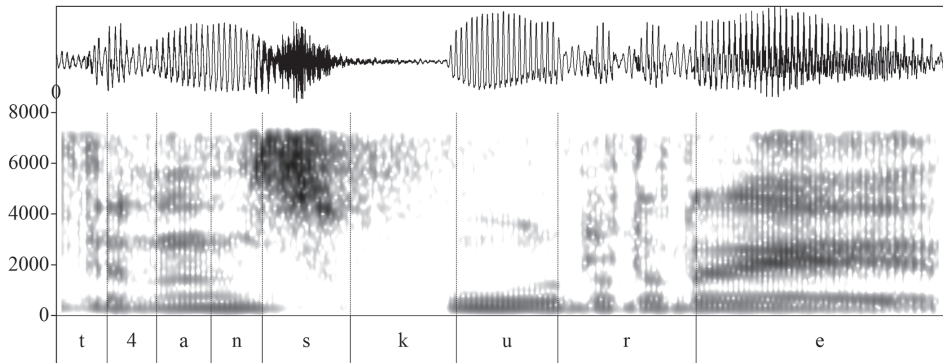


Figura 1: Una /s/ clasificada como retenida en la palabra “transcurre”

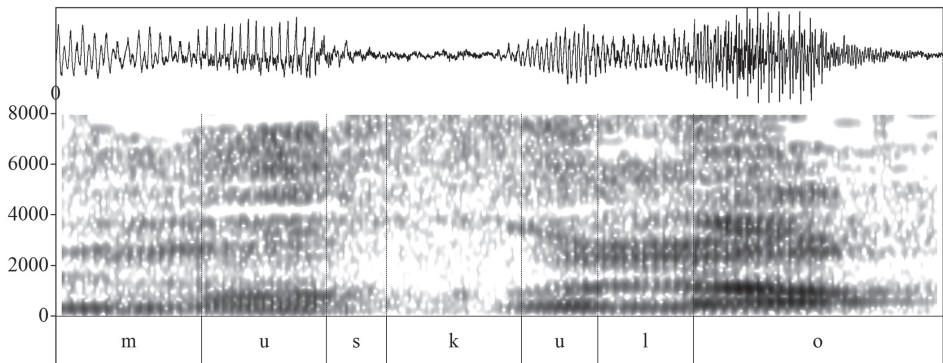


Figura 2: Una /s/ clasificada como debilitada en la palabra “músculo”

3.4. Análisis estadístico de las mediciones acústicas

En la Tabla 2 a continuación se presentan los principales resultados del análisis acústico realizado en el marco muestral ($n = 2977$).

Tabla 2: Distribución de las variables en el análisis acústico

	DUR [s]	COG [Hz]	UNVOI [%]
Valor mínimo	0.01124	1060	0.0000
1er cuartil	0.05000	2942	0.0000
Mediana	0.06979	5327	0.5556
Promedio	0.07647	4861	0.4766
3er cuartil	0.09329	6397	0.8000
Valor máximo	0.64000	9015	1.0000
Desviación típica	0.04078	1967	0.3585

Como veremos más adelante, los valores de mediana y promedio obtenidos en el análisis estadístico básico no son significativos para nuestro estudio.

Las Figuras 3, 4 y 5 nos proporcionan la distribución de probabilidad de la variable según cada uno de los tres parámetros acústicos estudiados.

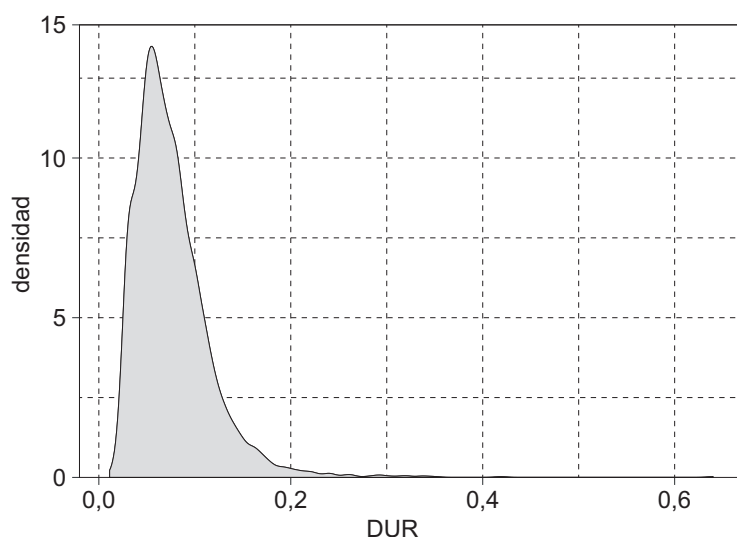


Figura 3: Distribución de la variable según DUR

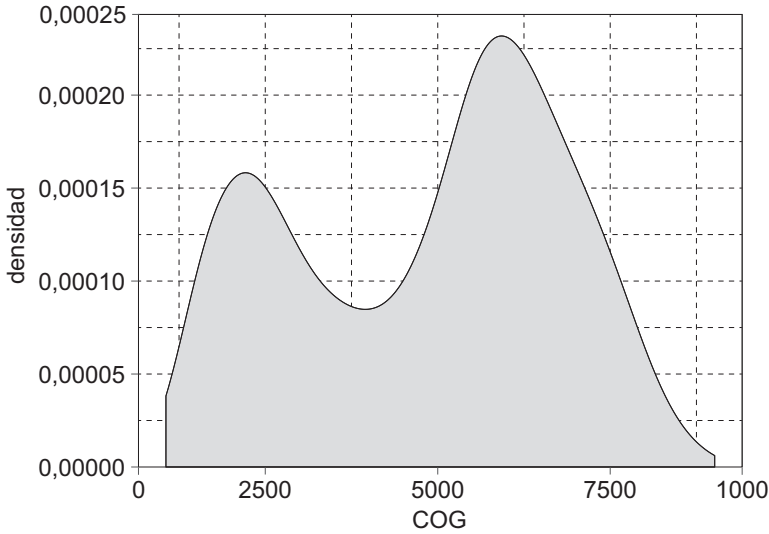


Figura 4: Distribución de la variable según COG

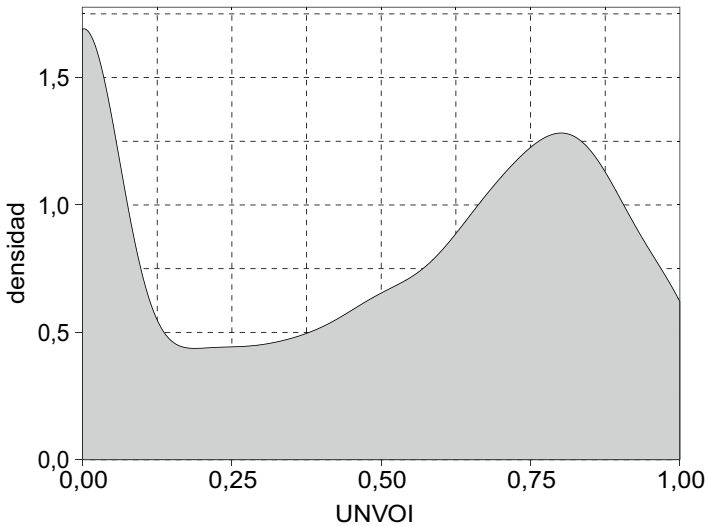


Figura 5: Distribución de la variable según UNVOI

Los gráficos de densidad con valores de la variable en el eje horizontal y densidad en el eje vertical muestran claramente dos picos en caso de COG y UNVOI, lo cual indica que los valores se concentran en dos grupos, mientras que los valores de DUR se acumulan casi en su totalidad entre 0.0 y 0.2 segundos. Esta observación nos lleva a suponer que, tras un cálculo adicional, podría establecerse una línea de división entre los alófonos retenidos y los debilitados, que visualizamos a continuación de manera provisional para el COG:

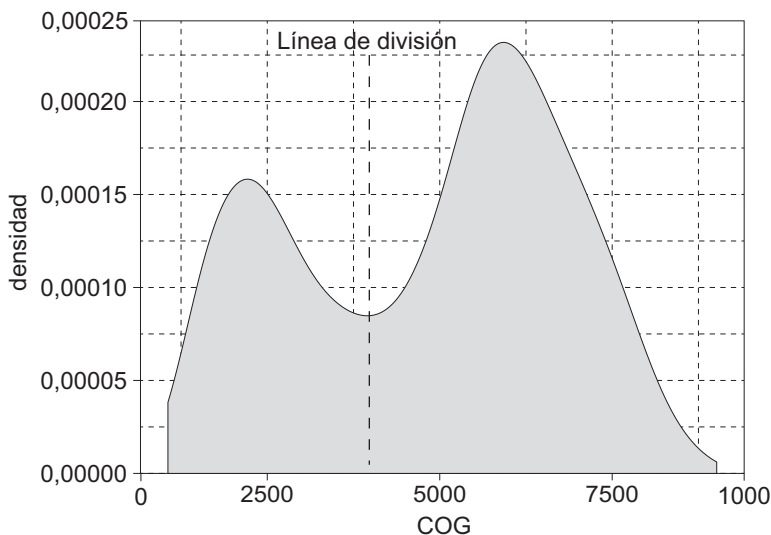


Figura 6: Simulación de una línea de división entre dos tipos de realizaciones

No obstante, el coeficiente de Pearson revela una correlación relativamente débil, aunque significativa, entre los parámetros.

En el eje diagonal aparecen los gráficos de distribución. En los paneles debajo del eje diagonal aparecen diagramas de dispersión. Por encima del eje diagonal aparecen los valores de la correlación de Pearson.

No existe una relación clara entre las observaciones y aparecen bastantes medidas incoherentes, en las cuales un valor bajo de COG va acompañado de un valor alto de UNVOI o al revés. Esto quiere decir que hay un número grande de realizaciones sordas de frecuencias bajas o sonorizadas de frecuencias altas.

3.5. Análisis estadístico de los resultados de la prueba perceptiva

Con el objetivo de encontrar una relación entre los parámetros acústicos y el tipo de realización de la consonante, hemos utilizado aprendizaje automatizado. Nos hemos planteado si es posible encontrar una correspondencia entre los valores de COG, UNVOI y DUR y el valor cualitativo de la fricativa. Este problema puede solucionarse sistemáticamente con el uso de métodos de clasificación, p. ej. regresión logística. Para realizar esta tarea, hemos usado un 30% de nuestro conjunto de datos anotado con el tipo de consonante (retenida/debilitada), donde las variables independientes eran COG, UNVOI y DUR.

Del total de las observaciones analizadas ($n = 898$), más de la mitad ($n = 579$, un 64%) corresponde a la realización retenida (sibilante), y 319 casos (un 36%) a la debilitada (aspirada).

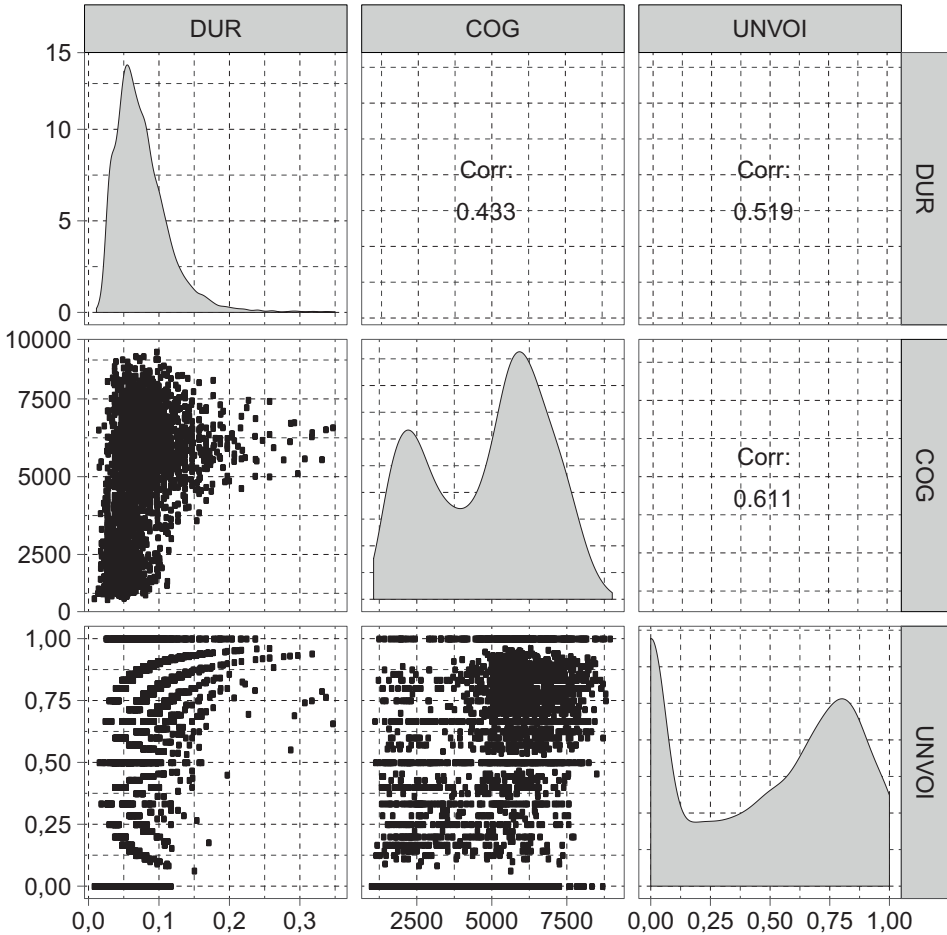


Figura 7: Correlación entre los parámetros acústicos (Pearson)

La Figura 8 muestra los resultados del análisis en forma de un diagrama de cajas. El gráfico presenta una clara división de los parámetros acústicos entre los dos grupos de consonantes clasificados como retenidas o debilitadas. Efectivamente, el valor medio de DUR en el grupo de retenidas es de 0.0913, mientras que en el grupo de debilitadas es de 0.0508 ($p < 0.00001$, prueba t de estudiante). De manera similar, el valor medio de UNVOI es igual a 0.640 para las retenidas y 0.183 para las debilitadas ($p < 0.00001$; prueba t de estudiante) y el valor medio de COG es de 6065 para retenidas y 2764 para debilitadas ($p < 0.00001$; prueba t de estudiante).

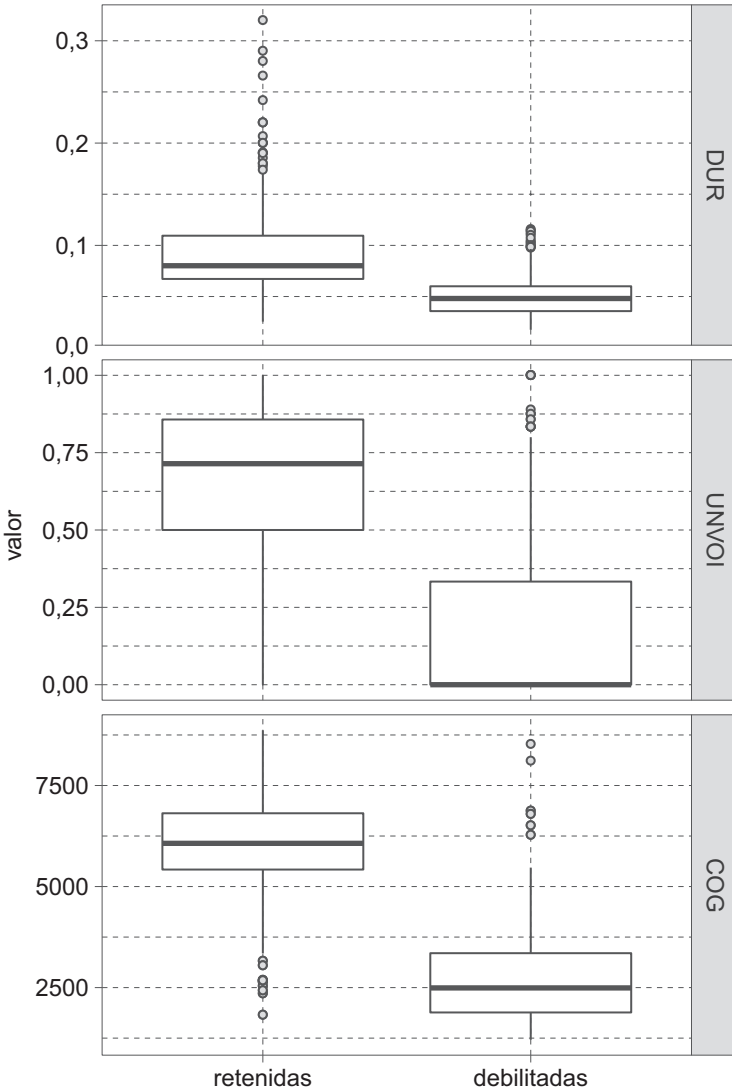


Figura 8: Relación entre los parámetros acústicos y el tipo de realización

En el gráfico se observa una clara división entre los dos tipos de realizaciones según cada uno de los tres parámetros con algunos valores atípicos. Esto se refleja en el valor de la sensibilidad (de las realizaciones de parámetros acústicos característicos para una /s/ debilitada, un 94% han sido clasificadas como debilitadas) y de la especificidad (de las realizaciones sibilantes, un 92% han sido clasificadas como tales). La exactitud del muestreo es del 93%. Esta herramienta sirve para encontrar la mejor manera de dibujar una línea entre los picos de los diagramas de densidad. La Tabla 3 compara los resultados

del clasificador logístico con una variable con los resultados del clasificador logístico con tres variables. Cada clasificador tiene el valor-p de la prueba de significancia de la exactitud a nivel de < 0.00001 .

Tabla 3: Comparación de los resultados del clasificador logístico con una variable con los del clasificador logístico con tres variables

	Clasificador logístico con tres variables	Clasificador logístico con una variable		
		DUR	COG	UNVOI
Exactitud	93%	82%	90%	82%
Exactitud 95% (intervalo de confianza)	(0.915, 0.949)	(0.789, 0.841)	(0.889, 0.924)	(0.794, 0.845)
Sensibilidad	94%	84%	92%	85%
Especificidad	92%	76%	89%	76%

Se advierte que el parámetro más exacto es el COG, mientras que la especificidad del DUR es muy baja, lo cual significa que muchas realizaciones retenidas (según los parámetros acústicos) fueron clasificadas como debilitadas en la prueba perceptiva.

4. Discusión y conclusiones

Los resultados presentados en los apartados anteriores no son conclusivos; sin embargo, aportan datos útiles para un estudio ampliado que sigue a este.

El estudio inicial presentado arriba es un paso en la elaboración de un algoritmo entrenado para la clasificación de las realizaciones de /s/ según sus tipos, basado en mediciones acústicas objetivas. En un corpus de datos grande, como CHARG, el tratamiento manual de los datos resulta imposible, por lo cual es inevitable elaborar un método automático. La clasificación de los alófonos en el corpus posibilitará un análisis de varios factores, tanto lingüísticos como extralingüísticos, que afectan la producción de la consonante en cuestión.

El valor alto de exactitud obtenido en el estudio presentado por medio de la regresión logística es un buen pronóstico para esta idea. No obstante, tienen que tomarse en cuenta algunas medidas.

Primero, los conjuntos de entrenamiento y prueba deberían ser seleccionados de una manera más heterogénea, abarcando una variedad de programas y estilos más amplia y evitando repeticiones de palabras. El estudio ha sido rea-

lizado en una parte del corpus. Se asume que la heterogeneidad de la muestra se garantizará cuando la prueba abarque todos los estilos provistos por CHARG.

Segundo, la evaluación debería ser realizada por un panel de al menos tres expertos para sacar conclusiones sólidas. Un algoritmo así tendría un carácter claramente aplicable para la evaluación de conjuntos grandes de datos en los que no se cuenta con las condiciones de laboratorio, como es el caso del habla pública.

Tercero, consideramos necesario evaluar otros parámetros acústicos, como p.ej. los cuatro momentos espectrales, los picos de energía o la amplitud absoluta. Estos parámetros indican el lugar de articulación de fricativas, por lo cual deben resultar especialmente útiles en el estudio del debilitamiento de /s/.

Un método de clasificación automática de los alófonos de /s/ será de gran utilidad para los conjuntos amplios de datos. La combinación de métodos estadísticos, pruebas perceptivas y análisis acústicos puede servir también para el estudio de otros sonidos de las lenguas. Siempre hay que advertir, sin embargo, que los métodos computacionales ven solo una parte de la realidad y, por lo tanto, deben ir acompañados por una evaluación humana, por más imperfecta que esta sea.

Referencias bibliográficas

BOERSMA Paul, WEENINK David

2017 "Praat: doing phonetics by computer" [software]. Versión 6.0.29, <<http://www.praat.org>>, 24 de mayo de 2017.

BROŚ Karolina

2015 *Survival of the Fittest: Fricative Lenition in English and Spanish from the Perspective of Optimality Theory*, Cambridge, Cambridge Scholars Publishing.

BROWN Earl K.

2011 "Paradigmatic Peer-Pressure: Word-Medial, Syllable-Initial /s/ Lenition in Dominican Spanish", en: Alvord S.M. (ed.), *Selected Proceedings of the 5th Conference on Laboratory Approaches to Romance Phonology*, Somerville, MA, Cascadilla Proceedings Project, pp. 46–58.

CANFIELD D. Lincoln

1988 *El español de América: Fonética*, Barcelona, Editorial Crítica.

CLEMENTS George N.

1990 "The role of the sonority cycle in core syllabification", en: Kingston J., Beckman M.E. (eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, Cambridge, Cambridge University Press, pp. 283–333.

ESTRELLA SANTOS Ana

2001 *El uso del verbo en el habla de Quito*, Quito, Abya-Yala.

2009 *El léxico de Pichincha y Guayas: un estudio comparativo*, Quito, Centro de Publicaciones PUCE.

FILE-MURIEL Richard J., BROWN Earl K.

2011 "The gradient nature of s-lenition in Caleño Spanish", *Language Variation and Change*, 23, pp. 223–243.

FLORES Tanya L.

- 2016 “Velar palatalization in Chilean public speech”, *Glossa: a journal of general linguistics*, 1 (6), pp. 1–17.

FLORES MEJÍA Esthela

- 2014 “Actitudes lingüísticas en Ecuador: Una tradición normativa que subsiste”, en: Chiquito A.B., Quesada Pacheco M.Á. (eds.), *Actitudes lingüísticas de los hispanohablantes hacia el idioma español y sus variantes*, Bergen, BeLLS, pp. 409–488.

GOLDMAN Jean-Philippe

- 2011 “EasyAlign: an automatic phonetic alignment tool under Praat”, en: *Proceedings of InterSpeech*, Firenze.

HABOUD Marleen

- 2004 “El español y las lenguas indígenas en el Ecuador hoy”, en: *III Congreso Internacional de la Lengua Española*, Rosario, Centro Virtual Cervantes.
- 2005 “Simultaneidad o perfectividad. El gerundio en el castellano andino”, *Revista UniverSOS*, Universidad de Valencia, 2, pp. 9–39.
- 2009 “Ecuador amazónico”, en: Sichra I. (ed.), *Atlas sociolingüístico de los pueblos indígenas de América Latina*, Vol. I, Quito, UNICEF/FUNPROEIB Andes, pp. 333–358.

INSTITUTO NACIONAL DE ESTADÍSTICAS Y CENSOS

- 2010a “Fascículo provincial Guayas” <<http://www.ecuadorencifras.gob.ec/wp-content/descargas/Manu-lateral/Resultados-provinciales/guayas.pdf>>, 18 de septiembre de 2019.
- 2010b “Fascículo provincial Pichincha” <<http://www.ecuadorencifras.gob.ec/wp-content/descargas/Manu-lateral/Resultados-provinciales/guayas.pdf>>, 18 de septiembre de 2019.

JONGMAN Allard, WAYLAND Ratre, WONG Serena

- 2000 “Acoustic characteristics of English fricatives”, *Journal of the Acoustical Society of America*, 108, pp. 1252–1263.

KLESSA Katarzyna, KARPINSKI Maciej, WAGNER Agnieszka

- 2013 “Annotation Pro – a new software tool for annotation of linguistic and paralinguistic features”, en: Hirst D., Bigi B. (eds.), *Proceedings of the Tools and Resources for the Analysis of Speech Prosody (TRASP) Workshop*, Aix-en-Provence, pp. 51–54.

LOC (Ley Orgánica de Comunicación),

- 2013 Registro Oficial Suplemento 22, 25/06/2013.

LÓPEZ GONZÁLEZ Antonio María

- 2001 *El lenguaje radiofónico en la ciudad de Almería. Análisis Sociolingüístico*, tesis doctoral, Almería, Universidad de Almería.

MARRERO Victoria

- 1990 “Estudio acústico de la aspiración en español”, *Revista de Filología Española*, LXX, pp. 345–397.

MARTÍNEZ CELDRÁN Eugenio, FERNÁNDEZ PLANAS Ana M.

- 2013 *Manual de fonética española. Coarticulaciones y sonidos*, segunda edición, Barcelona, Ariel Letras.

PALACIOS Azucena

- 2005 “La influencia del quichua en el español andino ecuatoriano”, en: Ferrero Pino C., Lasso-von Lan N. (eds.), *Varietades lingüísticas y lenguas en contacto en el mundo de habla hispana*, Bloomington, Books Library, pp. 44–52.
- 2006 “Cambios inducidos por contacto en el español de la sierra ecuatoriana: la simplificación de los sistemas pronominales (procesos de neutralización y elisión)”, en: *Huellas del contacto*, Vol. Monográfico de Tópicos del Seminario, Puebla (México), Benemérita Universidad Autónoma de Puebla, pp. 197–229.
- 2015 “De nuevo sobre la omisión de objeto directo en el español andino ecuatoriano”, *Círculo de Lingüística Aplicada a la Comunicación*, 61, pp. 104–130.

REAL ACADEMIA ESPAÑOLA

2011 *Nueva gramática de la lengua española. Fonética y fonología*, Vol. 3, Barcelona, Espasa Libros.

STRYCHARCZUK Patrycja, KOHLBERGER Martin

2016 “Resyllabification Reconsidered: On the Durational Properties of Word-Final /s/ in Spanish”, *Laboratory Phonology*, 7 (1), pp. 1–24.

STRYCHARCZUK Patrycja, VAN ‘T VEER Marijn, BRUIL Martine, LINKE Kathrin

2014 “Phonetic evidence on phonology–morphosyntax interactions: Sibilant voicing in Quito Spanish”, *Journal of Linguistics*, 50, pp. 403–452.

TEJADA GIRÁLDEZ María de la Sierra

2015 *Convergencia y divergencia entre comunidades de habla: a propósito de la /s/ implosiva. Contribución al estudio de los patrones sociolingüísticos del español de Granada*, tesis doctoral, Granada, Universidad de Granada.

TOSCANO MATEUS Humberto

1953 *El español en el Ecuador*, Madrid.

WALKER Abby, GARCÍA Christina, CORTÉS Yomi, CAMPBELL-KIBLER Kathryn

2014 “Comparing social meanings across listener and speaker groups: The indexical field of Spanish /s/”, *Language, Variation and Change*, 26, pp. 169–189.

Sitios de Internet

Earl K. Brown (Página personal), Department of Modern Languages, Kansas State University, <www-personal.ksu.edu/~ekbrown>, 12 de marzo de 2019.

Toward an automatized classification of /s/-allophones in Guayaquil Spanish

Keywords: /s/-weakening — machine learning approach — Guayaquil Spanish — speech processing — acoustic phonetics.

Abstract

The paper aims to generate a way of automatized classification of /s/-allophones by finding a correspondence between their acoustic parameters and the qualitative type of the realization. The dataset used for the study comes from CHARG (Corpus de Habla Radiofónica de Guayaquil). The parameters considered for the acoustic description of /s/ are duration (DUR), center of gravity (COG) and percentage of voicelessness (UNVOI), since they are reported to indicate the lenition of the phoneme. To perform the task, we retrieved the acoustic parameters’ values. Next, we excerpted randomly 30% of the tokens and classified them perceptually as either retained or weakened. Statistical analysis revealed a weak, despite significant, correlation between the parameters, but a satisfactory correlation between DUR/COG/UNVOI and the results of the perceptual test, which is a good prediction for the elaboration of a trained algorithm for classifying the fricatives in big datasets of non-laboratory speech samples.

Fecha de recepción: 12 de noviembre de 2019

Fecha de aceptación: 16 de abril de 2019