

IMPORTANCIA ECONÓMICA DE LAS CARACTERÍSTICAS FONÉTICAS DEL IDIOMA ESPAÑOL Y SUS VARIEDADES REGIONALES

Germán Coloma (Universidad del CEMA, Buenos Aires, Argentina)

Resumen

El presente trabajo busca cuantificar económicamente la importancia de cinco características fonéticas (seseo, yeísmo, aspiración del fonema /s/, rehilamiento del fonema /z/ y aspiración del fonema /x/), cuya presencia o ausencia sirve para distinguir entre diez variedades regionales del idioma español. Para ello se efectúa un análisis de la extensión de dichas características y se estiman las que corresponden al “hablante promedio” de español, que resultan ser idénticas a las del “español mexicano-centroamericano”. A través de estimaciones de ingreso per cápita, sin embargo, se llega a la conclusión de que la característica asociada con un mayor ingreso es la ausencia de seseo (típica del español peninsular norteño). Por último, mediante un análisis de regresión, se verifica que dicha característica está asociada con un ingreso anual medio por habitante que es de U\$S 20.000 más que el de los hablantes seseantes. Dicho análisis nos muestra también que las otras dos características estadísticamente significativas como señales de un ingreso más elevado son la presencia de yeísmo y la ausencia de aspiración del fonema /x/.

Palabras clave: características fonéticas, variedades regionales del español, hablante promedio, ingreso per cápita, regresión.

Clasificación del JEL: Z13, A12.

Abstract

This paper tries to quantify economically the importance of five phonetic characteristics (*seseo*, *yeísmo*, /s/-aspiration, *rehilamiento*, and /x/-aspiration) whose presence or absence allows to distinguish among ten regional varieties of Spanish. To do that we perform an analysis of the extension of those characteristics and we estimate which of them correspond to the “average speaker” of Spanish, which end up being identical to the ones of the so-called “Mexican-Central American Spanish”. Through estimations of per capita income, however, we find that the characteristic associated with a higher income level is the absence of *seseo*, typical of the Northern Peninsular Spanish. Finally, through a regression analysis, we find that such characteristic is associated with an average annual per capita income which is US\$ 20,000 higher than the one of Spanish speakers that possess *seseo*. We also find that the other additional characteristics which are statistically significant as signals of a higher income are the presence of *yeísmo* and the absence of /x/-aspiration.

Keywords: phonetic characteristics, Spanish regional varieties, average speaker, per capita income, regression.

JEL Classification: Z13, A12.

IMPORTANCIA ECONÓMICA DE LAS CARACTERÍSTICAS FONÉTICAS DEL IDIOMA ESPAÑOL Y SUS VARIEDADES REGIONALES

Germán Coloma (Universidad del CEMA, Buenos Aires, Argentina)

El objetivo de este artículo es analizar las principales características fonéticas que diferencian a las distintas variedades regionales del idioma español y cuantificar su importancia relativa en términos económicos. Para ello, el trabajo se divide en una primera sección que describe los principales sonidos y fonemas que se utilizan para hablar la lengua española, una segunda sección que describe los principales sistemas fonológicos vigentes en las zonas en las que se habla el idioma español, una tercera sección en la cual se hace un análisis cuantitativo de la importancia demográfica y económica de las distintas características encontradas, una cuarta sección en la cual se relaciona a las características fonéticas con el ingreso promedio de los hablantes de español, y una última sección de conclusiones.

1. Sonidos y fonemas del idioma español

Los sonidos que se utilizan para hablar el idioma español, al igual que los de la mayoría de las lenguas de origen europeo, se pueden clasificar de manera bastante tajante en vocales y consonantes. Los sonidos vocálicos del español son cinco, y se representan a través de los signos fonéticos [a], [e], [i], [o] y [u]. Dichos sonidos se diferencian entre sí por su punto de articulación, que puede ser más o menos abierto o cerrado, y más o menos anterior o posterior (ver cuadro 1).

Cuadro 1: Sonidos vocálicos del español

Tipo de vocal	Anterior	Central	Posterior
Cerrada	i		u
Media	e		o
Abierta		a	

Cada uno de estos cinco sonidos vocálicos tiene a su vez características adicionales que lo identifican. Los sonidos [a], [e] e [i], por ejemplo, son vocales no labializadas; en tanto que los sonidos [o] y [u] son vocales labializadas. Cada uno de ellos puede presentar ciertas variaciones en cuanto a su grado de apertura y posicionamiento, pero en todas las posiciones en las que aparecen dentro del idioma español mantienen sus características básicas y su capacidad de oponerse unos a otros. Esto vale tanto para los casos de sílabas tónicas como átonas, y para los casos en los cuales los sonidos vocálicos aparecen formando diptongos o triptongos.

En cuanto a los sonidos consonánticos del español, los mismos se diferencian entre sí según su punto de articulación y también según su modo de articulación. Según su punto de articulación, dichos sonidos pueden ser labiales, dentales, alveolares, palatales, velares o glotales. Según su modo de articulación, la principal distinción que puede hacerse entre ellos es la que los divide en consonantes obstruyentes y sonantes, entrando dentro del primer grupo los sonidos oclusivos y fricativos (que a su vez pueden ser sordos o sonoros) y dentro del segundo las consonantes nasales, laterales, róticas y semivocales (ver cuadro 2)¹.

A efectos de interpretar el cuadro 2, cabe aclarar algunas convenciones que se han utilizado. En primer lugar, en la columna de sonidos dentoalveolares se han incluido algunos sonidos que son propiamente dentales (por ejemplo, [θ]), otros que son alveolares (por ejemplo, [r] y [r]), y otros que pueden ser dentales o alveolares. Lo mismo puede decirse de la columna de sonidos alveopalatales, en la cual aparece el sonido alveolar [s], los sonidos palatales [ɲ], [ʎ] y [j], y los sonidos postalveolares [ʒ] y [tʃ]. Este último sonido no es en

¹ Para una descripción detallada de todos estos conceptos fonéticos, véase Martínez y Fernández (2007).

realidad oclusivo sino africado, es decir, es un sonido que presenta un elemento oclusivo seguido de un elemento fricativo. En segundo lugar, debe aclararse que los sonidos obstruyentes sonoros ([b], [d], [ʒ] y [g]) tienen una realización oclusiva o africada ([b], [d], [dʒ] y [g]) y una realización alternativa fricativa ([β], [ð], [ʒ] y [ɣ]), que hemos omitido en el cuadro por cuestiones de simplicidad. En tercer lugar, corresponde mencionar que, para el caso de las consonantes róticas, el sonido [r] representa a una consonante vibrante simple, en tanto que el sonido [r̄] corresponde a una consonante vibrante múltiple. Por último, debemos mencionar que la semivocal velar [w] es en rigor labiovelar, puesto que se articula utilizando tanto el velo del paladar como los labios².

Cuadro 2: Sonidos consonánticos del español

Tipo de consonante	Labial	Dentoalveolar	Alveopalatal	Velar	Glotal
Oclusiva sorda	p	t	tʃ	k	
Fricativa sorda	f	θ	s	x	h
Obstruyente sonora	b	d	ʒ	g	
Nasal	m	n	ɲ		
Lateral		l	ʎ		
Rótica		r	r̄		
Semivocal			j	w	

En base a los 27 sonidos básicos descritos (cinco vocales más veintidós consonantes), el idioma español articula los distintos fonemas que constituyen su sistema fonológico. Para muchos de ellos existe una única realización (es decir, se pronuncian utilizando un único sonido), pero en algunos casos los hablantes utilizan más de un sonido para articular un único fonema, apareciendo en dichas circunstancias lo que en fonética se denomina “alófonos”. Existen además en español ciertas variaciones regionales, que hacen que hablantes de distintas regiones utilicen distintos sonidos para articular el mismo fonema³.

El número de fonemas del español es también distinto según la variedad regional que se analice. Si bien la mayoría de los textos sobre fonología española describen como variedad principal un tipo de lengua que emplea 24 fonemas distintos, la mayoría de los hablantes de español utilizan variedades que sólo usan 22 ó a lo sumo 23 fonemas. Al igual que en el caso de los sonidos que los forman, los fonemas también pueden clasificarse en principio en fonemas vocálicos y consonánticos. Los fonemas vocálicos del español son cinco: /a/, /e/, /i/, /o/ y /u/. Los fonemas /a/, /e/ y /o/ se articulan a través de los sonidos vocálicos [a], [e] y [o]. Los fonemas /i/ y /u/, en cambio, se articulan a través de los sonidos vocálicos [i] y [u], salvo en los casos en los cuales aparecen formando diptongos. En tales casos, el fonema /i/ se articula a través del sonido semivocal palatal [j], y el fonema /u/ se articula a través del sonido semivocal labiovelar [w].

En lo que se refiere a los fonemas puramente consonánticos, los 17 fonemas obligatorios del español son /b/, /d/, /f/, /g/, /k/, /l/, /m/, /n/, /ɲ/, /p/, /r/, /r̄/, /s/, /t/, /tʃ/, /x/ y /ʒ/. Existen además otros dos fonemas optativos (/ʎ/ y /θ/), que aparecen en algunas variedades regionales del idioma español y no en otras. En aquellas variedades en las cuales los mismos no se utilizan, se reemplaza al fonema /ʎ/ por el fonema /ʒ/ (yeísmo) y al fonema /θ/ por el fonema /s/ (seseo).

La mayoría de los fonemas consonánticos del idioma español se articula a través de un único sonido básico, siendo las principales excepciones los fonemas /s/, /x/ y /ʒ/. El fonema /s/ tiene en algunas variedades regionales al sonido [h] como alófono, en especial cuando el mismo aparece en posición final de sílaba y precediendo a una consonante (a

² Agradezco los comentarios de Laura Colantoni respecto de estos temas.

³ Para un análisis más detallado de estos temas véase Coloma (2010), del cual hemos extraído buena parte de los contenidos de las secciones 1 y 2 del presente trabajo.

este fenómeno se lo denomina comúnmente “aspiración” del fonema /s/. El sonido [h] también se usa en lugar del sonido [x] en algunas variedades regionales para articular el fonema /x/ (aspiración de /x/). En lo que respecta al fonema /ʒ/, la mayoría de las variedades regionales del idioma español lo articula utilizando el sonido [j], salvo cuando el mismo se encuentra después de una consonante nasal o lateral (en cuyo caso se utiliza el sonido [ʒ] o [dʒ]). Otras variedades regionales, sin embargo, utilizan el sonido [ʒ] como realización principal del fonema /ʒ/ en todas las posiciones, y a este fenómeno se lo conoce como “rehilamiento”⁴.

2. Variedades regionales del idioma español

La literatura sobre dialectología del español ha elaborado diferentes propuestas para subdividir las variedades regionales de la lengua castellana. Piñeros (2008), por ejemplo, distingue tres grandes geoelectos en los que puede dividirse el mundo hispanohablante, a los que denomina “peninsular norteño” (que abarca el centro y norte de España), “terrabajense” (que abarca el sur de España y numerosas zonas de América) y “terraltense” (que abarca el resto de las zonas del continente americano en las cuales se habla español como primera lengua). Usando una división geográfica más detallada, Moreno Fernández y Otero (2006) distinguen ocho áreas en las que subdividen a la lengua española, que son el área mexicano-centroamericana, el área caribeña, el área andina, el área chilena, el área rioplatense, el área andaluza, el área canaria y el área castellana. Esta última clasificación encaja exactamente dentro de la anterior, ya que las variedades mexicano-centroamericana y andina serían de tipo terraltense, las variedades caribeña, chilena, rioplatense, andaluza y canaria serían de tipo terrabajense, y la variedad castellana sería de tipo peninsular norteño.

Si empleamos una caracterización estrictamente fonológica de las principales diferencias que aparecen entre las distintas variedades del idioma español, sin embargo, las zonas en las cuales corresponde dividir al mundo hispanohablante son un poco diferentes de las reseñadas por estos autores. La zona peninsular norteña, por ejemplo, se caracteriza básicamente por ser la única parte del mundo hispánico que no es seseante (o sea, que distingue entre el fonema /s/ y el fonema /θ/), y dentro de ella resulta posible distinguir un área yeísta y un área no yeísta. A la variedad de español que se habla en la primera de tales áreas la denominaremos “español peninsular estándar”, en tanto que a la variedad no yeísta del español peninsular norteño la denominaremos “castellano tradicional”⁵.

La división fonética entre la zona terraltense y la zona terrabajense (ambas seseantes), tiene en cambio que ver con la presencia o ausencia de aspiración generalizada del fonema /s/ en posición preconsonántica. Dicha aspiración es característica de las variedades terrabajenses y no de las terraltenses. En cada una de dichas zonas, sin embargo, conviven dialectos yeístas y no yeístas. La variedad terrabajense no yeísta es la que se habla en Paraguay y en algunas provincias argentinas adyacentes a dicho país (español paraguayo), en tanto que la variedad terraltense no yeísta se habla en algunas zonas de Colombia, Perú y Bolivia (español andino tradicional).

En la zona terraltense yeísta, por su parte, la principal subdivisión fonética adicional que puede llevarse a cabo tiene que ver con la presencia o ausencia de aspiración del fonema /x/. Este fenómeno, que hace que dicho fonema se articule a través del sonido [h] en vez del sonido [x], es un rasgo que no es común en el área que Moreno Fernández y Otero denominan “mexicano-centroamericana” (que abarca México, Guatemala, El Salvador, Nicaragua, Honduras y Costa Rica), y es en cambio muy frecuente en el área andina (Ecuador, Perú, Bolivia y parte de Colombia)⁶. Al dialecto terraltense que no

⁴ De hecho, el rehilamiento del fonema /ʒ/ implica que el mismo puede ser pronunciado como [ʒ], como [j] o como [dʒ], pero nunca como un sonido semivocal [j]. Sobre este tema, véase Colantoni (2008).

⁵ Esta forma de identificar las variedades regionales del español es semejante a la que utiliza la literatura sobre reconocimiento automático de dialectos. Véase, por ejemplo, Caballero, Moreno y Nogueiras (2009).

⁶ Sobre este punto, véase Avila (2003).

presenta aspiración de /x/ lo llamaremos por lo tanto “español mexicano-centroamericano”, en tanto que al dialecto terraltense yeísta que sí presenta dicha aspiración lo llamaremos “español andino moderno”.

La aspiración del fonema /x/ es un rasgo que también aparece en las variedades del español terrabajense, y que permite diferenciar por un lado entre el español rioplatense, chileno y paraguayo (que no aspiran el fonema /x/) y por otro entre el español caribeño y el andaluz-canario (que sí aspiran el fonema /x/). En lo que se refiere al rehilamiento del fonema /ʒ/ en todas las posiciones, el mismo es un rasgo que ha sido señalado por Piñeros (2008) como característico del español del sur de España (andaluz y canario), y sirve por lo tanto para diferenciar a dicha variedad del español caribeño (que es el que se habla en Cuba, República Dominicana, Puerto Rico, Panamá, Venezuela y parte de Colombia). Dicho rehilamiento del fonema /ʒ/ es también característico del español rioplatense, y sirve por lo tanto para diferenciar entre este último dialecto y el español chileno. El español paraguayo, por su parte, también se caracteriza por rehilar el fonema /ʒ/, pero su gran diferencia con el español rioplatense es que se trata de una variedad no yeísta (es decir, que distingue entre el fonema /ʒ/ y el fonema /ʎ/).

Con estos elementos podemos entonces completar la caracterización fonética de las diez variedades regionales del idioma español que hemos identificado, que son el castellano tradicional (CT), el español peninsular estándar (PS), el español mexicano-centroamericano (MC), el español andino tradicional (AT), el español andino moderno (AM), el español chileno (CH), el español paraguayo (PA), el español rioplatense (RP), el español caribeño (CB) y el español andaluz-canario (AC). Dicha caracterización es la que aparece en el cuadro 3.

Cuadro 3: Diferencias fonéticas entre las variedades regionales del español

Característica	CT	PS	MC	AT	AM	CH	PA	RP	CB	AC
Seseo	No	No	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Yeísmo	No	Sí	Sí	No	Sí	Sí	No	Sí	Sí	Sí
Aspiración de /s/	No	No	No	No	No	Sí	Sí	Sí	Sí	Sí
Rehilamiento de /ʒ/	No	No	No	No	No	No	Sí	Sí	No	Sí
Aspiración de /x/	No	No	No	Sí	Sí	No	No	No	Sí	Sí

Tal como puede observarse, ninguna de estas diez variedades regionales es idéntica a otra en las cinco características fonéticas que hemos descripto. El orden en el cual las mismas aparecen en el cuadro 3 va desde las variedades más conservadoras hasta las menos conservadoras, notándose que en un extremo aparece el castellano tradicional (caracterizado por no poseer ni seseo, ni yeísmo, ni aspiración del fonema /s/, ni rehilamiento del fonema /ʒ/, ni aspiración del fonema /x/) y en el otro aparece el español andaluz-canario (caracterizado por poseer todas esas características fonéticas).

Asimismo, el cuadro 3 nos indica que la única de las diez variedades descriptas que utiliza 24 fonemas es el castellano tradicional (que presenta al mismo tiempo ausencia de seseo y de yeísmo), en tanto que con 23 fonemas aparecen el español peninsular estándar (yeísta pero no seseante), y las variedades andina tradicional y paraguaya (seseantes pero no yeístas). El resto de las variedades reseñadas utiliza 22 fonemas, por ser al mismo tiempo yeístas y seseantes.

3. Importancia demográfica y económica de las características del español

La tipificación de las distintas variedades del idioma español, de acuerdo con las características fonéticas que poseen, nos posibilitan efectuar un análisis cuantitativo de la importancia relativa de dichas características dentro del universo de hablantes del idioma español. Para ello es necesario cuantificar primero de manera aproximada el número de hablantes de las distintas variedades de español descriptas. Si queremos tener una idea respecto de la importancia económica de dichas variedades, resulta también útil hacer una

cuantificación del producto bruto generado por dichos hablantes.

Cuadro 4: Habitantes y producto bruto por país hispanohablante (2008)

País	Habitantes		Producto bruto	
	Millones	Porcentaje	Mil mill US\$	Porcentaje
España	45,56	11,14%	1.442,94	26,00%
México	106,35	26,00%	1.549,49	27,92%
Guatemala	13,69	3,35%	65,15	1,17%
Nicaragua	5,67	1,39%	15,24	0,27%
El Salvador	6,13	1,50%	41,71	0,75%
Costa Rica	4,52	1,10%	50,76	0,91%
Honduras	7,32	1,79%	28,78	0,52%
Panamá	3,40	0,83%	42,48	0,77%
Cuba	11,21	2,74%	110,80	2,00%
Rep Dominicana	9,95	2,43%	80,87	1,46%
Puerto Rico	3,95	0,97%	86,50	1,56%
Venezuela	27,94	6,83%	358,07	6,45%
Colombia	45,01	11,01%	395,99	7,14%
Ecuador	13,48	3,30%	108,04	1,95%
Peru	28,84	7,05%	245,36	4,42%
Bolivia	9,69	2,37%	41,46	0,75%
Chile	16,80	4,11%	242,58	4,37%
Argentina	39,88	9,75%	570,85	10,29%
Uruguay	3,33	0,82%	42,49	0,77%
Paraguay	6,24	1,53%	29,35	0,53%
Total	408,96	100,00%	5.548,89	100,00%

Un primer paso para llevar a cabo estas tareas consiste en cuantificar la población y el producto bruto de los países hispanohablantes. Para ello hemos tomado como base información del Banco Mundial (2008). Dicha información tiene la ventaja de que el producto bruto aparece medido utilizando unidades monetarias comparables, que son dólares estadounidenses ajustados por el concepto de “paridad del poder adquisitivo” (cuadro 4). En el cuadro aparecen sin embargo dos casos para los cuales no pudimos obtener información del Banco Mundial en lo que se refiere al producto bruto, que son Cuba y Puerto Rico. Para tales casos recurrimos a una fuente alternativa (CIA, 2009), que utiliza la misma metodología de cálculo del producto bruto.

En base a las cifras del cuadro 4 resulta posible estimar el número de hablantes (y el producto bruto asociado con dichos hablantes) que corresponde a cada una de las variedades regionales del idioma español que hemos identificado en la sección anterior. Para ello es necesario identificar primero el tipo de español hablado en cada país y, en el caso de países en los que se habla más de una variedad, estimar qué porcentaje de su población y de su producto bruto corresponde a cada una de los dialectos identificados. Para los casos del español mexicano-centroamericano y chileno, el cálculo es relativamente simple, ya que el segundo de ellos se habla solamente en Chile (y su número de hablantes y el producto bruto asociado a ellos pueden considerarse como sinónimos de los correspondientes a la República de Chile), y las cifras correspondientes al primero de ellos resultan simplemente de agregar los datos de México, Guatemala, Nicaragua, El Salvador, Costa Rica y Honduras.

Para el caso de España, en cambio, hemos dividido a su población y a su producto bruto entre las tres variedades idiomáticas individualizadas (castellano tradicional, peninsular estándar y andaluz-canario) usando como base los datos de población y producto bruto regional que aparecen en Eurostat (2001). En base a tales cifras calculamos los porcentajes de la población y el producto correspondiente a las tres áreas idiomáticas, haciendo el supuesto de que el castellano tradicional corresponde a la región de Castilla y

León, de que el español andaluz-canario corresponde a las regiones de Andalucía, Ceuta y Melilla, y Canarias, y de que en el resto de las regiones predomina el español peninsular estándar.

Otro país para el cual fue necesario dividir la población entre tres dialectos es Colombia, ya que en él coexisten el español andino tradicional, el español andino moderno y el español caribeño. El supuesto que utilizamos aquí fue que la primera de tales variedades corresponde a los departamentos de Antioquia, Boyacá, Nariño, Santander y Norte de Santander; que el español caribeño corresponde en cambio a los departamentos de Atlántico, Bolívar, César, Chocó, Córdoba, La Guajira, Magdalena, San Andrés y Valle; y que en el resto de los departamentos predomina el español andino moderno⁷. En Perú y en Bolivia, por su parte, dividimos a la población y al producto bruto en dos áreas, representativas del español andino tradicional y del español andino moderno. En Perú consideramos como área andina moderna a los departamentos de Ancash, Arequipa, Cajamarca, Ica, La Libertad, Lambayeque, Lima y Callao, Moquegua, Piura, Tacna y Tumbes; y al resto del país como hablante de español andino tradicional. En Bolivia, en cambio, consideramos andino tradicional al español hablado en Cochabamba, La Paz, Beni y Pando, y andino moderno al del resto de los departamentos bolivianos⁸.

Cuadro 5: Habitantes y producto bruto por variedad de español (2008)

Variedad de español	Hablantes		Producto bruto	
	Millones	Porcentaje	Mil mill US\$	Porcentaje
Castellano tradicional	2,89	0,71%	83,66	1,51%
Peninsular estándar	32,35	7,91%	1.111,79	20,04%
Andaluz-canario	10,31	2,52%	247,49	4,46%
Mexicano-centroamericano	143,68	35,13%	1.751,12	31,56%
Caribeño	66,45	16,25%	740,64	13,35%
Andino moderno	60,61	14,82%	557,50	10,05%
Andino tradicional	26,41	6,46%	171,43	3,09%
Chileno	16,80	4,11%	242,58	4,37%
Rioplataense	40,58	9,92%	596,57	10,75%
Paraguayo	8,88	2,17%	46,11	0,83%
Total	408,96	100,00%	5.548,89	100,00%

En la Argentina, por último, hemos considerado que la población y el producto bruto de las provincias de Formosa, Corrientes y Misiones corresponden al área de español paraguayo, y que el resto del país corresponde al área de español rioplatense⁹. La primera de tales zonas fue por lo tanto agregada a los datos de Paraguay, y la segunda a los datos de Uruguay. Para estimar la población y el producto asociados con el español caribeño, por su parte, sumamos los datos correspondientes a Cuba, República Dominicana, Puerto Rico, Panamá, Venezuela y los departamentos caribeños de Colombia; en tanto que para el español andino moderno agregamos la porciones ya definidas de Colombia, Perú y Bolivia junto con la totalidad de la República de Ecuador. Los resultados de todas estas agregaciones son los que aparecen en el cuadro 5, en el cual los totales, tal como puede verse, coinciden con los del cuadro 4.

Una simplificación a tener en cuenta en nuestros análisis cuantitativos es que, por más que en varios países hispanohablantes el español coexiste con otros idiomas (catalán, gallego, vasco, guaraní, quichua, etc.), en nuestro análisis hemos trabajado usando la idea de que los habitantes de los veinte países analizados hablaban también español. Del mismo

⁷ Los datos para realizar la división porcentual de la población y del producto bruto fueron en este caso tomados de DANE (2007).

⁸ Para el caso de Perú, la fuente de datos para dividir a la población y al producto bruto en regiones es INEI (2007), y en el de Bolivia es INE (1990; 2008).

⁹ Para el caso de la Argentina, la fuente de datos en la cual se basa la división es un informe de Abeceb (2008), que a su vez utiliza como base información del INDEC.

modo, hemos omitido del análisis a aquellos hablantes que viven en otros países, por ser muy difícil identificar qué variedad de español utilizan.

En base a las cifras que aparecen en el cuadro 5 y a la relación entre variedades regionales y características fonéticas que hemos resumido en el cuadro 3, resulta posible elaborar un nuevo cuadro en el que se relacionen dichas características fonéticas con los datos de población y producto bruto. Eso es lo que aparece en el cuadro 6, en el cual hemos incluido también columnas de porcentaje referidas a la importancia relativa de dichas características en términos demográficos y económicos.

Cuadro 6: Habitantes y producto bruto por característica fonética (2008)

Característica	Hablantes		PBI	
	Millones	Porcentaje	Mil mill US\$	Porcentaje
Seseo	373.718	91,38%	4.353.438	78,46%
Yeísmo	370.783	90,66%	5.247.688	94,57%
Aspiración de /s/	143.028	34,97%	1.873.391	33,76%
Rehilamiento de /ʒ/	59.769	14,61%	890.168	16,04%
Aspiración de /x/	163.784	40,05%	1.717.054	30,94%

De las cifras que aparecen en el cuadro 6 puede inferirse que algunas de las características fonéticas descritas son muy mayoritarias entre los hablantes del idioma español (como por ejemplo el seseo y el yeísmo, que son compartidos por más del 90% de los hispanohablantes), en tanto que otras son relativamente minoritarias (por ejemplo, el rehilamiento del fonema /ʒ/ y la aspiración de los fonemas /s/ y /x/). También puede verse que la importancia del seseo es bastante diferente si la medimos en términos demográficos (porcentaje de la población) y en términos económicos (porcentaje del producto bruto). Esto se debe a que la ausencia de seseo es privativa de la zona norte de España, que es justamente el área en la cual el ingreso medio por habitante es mayor dentro del mundo hispanohablante. Esto hace que, si bien la distinción entre /s/ y /θ/ sólo sea relevante para un 8,62% de los hablantes de español, dicho porcentaje de la población genere un 21,54% del producto bruto total de los países hispanohablantes.

Las cifras que aparecen en el cuadro 6 son también útiles para caracterizar un “hablante promedio del idioma español”. Dicho hablante sería alguien que compartiera las características fonéticas mayoritarias en términos de porcentajes, y nuestras cifras nos indican que se trata de una persona que posee al mismo tiempo seseo y yeísmo, pero que no aspira los fonemas /s/ y /x/ ni rehila el fonema /ʒ/. Dichas características mayoritarias no se modifican en términos cualitativos si medimos los porcentajes en términos de población o en términos de producto bruto, y tienen la particularidad de coincidir exactamente con las características que presenta una variedad regional determinada del idioma castellano, que es el español mexicano-centroamericano (ver cuadro 3).

4. Características fonéticas e ingreso por habitante

El análisis cuantitativo llevado a cabo en la sección anterior culminó con el resultado de que el español promedio que se habla en el mundo puede aproximarse en cuanto a sus características fonéticas a través del dialecto que se usa mayoritariamente en México y América Central. Este es un resultado esperado porque dicha zona concentra un porcentaje bastante elevado de la población hispanohablante total, pero en cierto modo contrasta con el hecho de que la mayor parte de la literatura sobre fonética española se ha concentrado en analizar otras variedades del idioma español, en especial el castellano tradicional y en menor medida el andaluz-canario y el rioplatense.

Una posible explicación socio-económica a dicho interés puede tener que ver con el hecho de que, si bien dichas variedades presentan en algún sentido características menos mayoritarias que el español mexicano-centroamericano, están asociadas con mayores niveles de ingreso por habitante. Al respecto resulta útil observar las cifras del cuadro 7, en

el cual hemos calculado el producto bruto interno per cápita (es decir, el cociente entre producto bruto y población) para cada una de las variedades regionales del español, y cuánto mayor o menor es dicho PBI per cápita en relación al promedio de toda el área hispanohablante.

Cuadro 7: PBI per cápita por variedad regional (2008)

Variedad de español	U\$S/Año	% s/Promedio
Castellano tradicional	28.966	113,48%
Peninsular estándar	34.363	153,26%
Andaluz-canario	23.995	76,85%
Mexicano-centroamericano	12.188	-10,17%
Caribeño	11.145	-17,86%
Andino moderno	9.199	-32,20%
Andino tradicional	6.491	-52,16%
Chileno	14.436	6,40%
Rioplataense	14.702	8,36%
Paraguayo	5.193	-61,72%
Promedio	13.568	0,00%

Las cifras del cuadro 7 nos muestran que el ingreso por habitante promedio de toda la zona hispanohablante es de U\$S 13.568 anuales, pero que dicho valor, por ejemplo, es un 153,26% mayor en la zona del español peninsular estándar y un 61,72% menor en la zona del español paraguayo. En el cuadro se observa también que las únicas variedades regionales americanas que están asociadas con un ingreso por habitante mayor al promedio son el español rioplataense y el español chileno, y que en cambio todas las variedades del español peninsular están asociadas con niveles de PBI per cápita mucho mayores que el promedio.

Estas diferencias de ingreso por habitante entre variedades regionales se extienden a las características fonéticas que distinguen a dichas variedades. En el cuadro 8 hemos calculado el PBI per cápita promedio de los hablantes que poseen y que no poseen cada una de las cinco características fonéticas analizadas en el presente trabajo, y vemos que también aquí las diferencias son notables. En promedio, por ejemplo, los hablantes no seseantes tienen un ingreso por habitante de U\$S 33.921 anuales, y eso es un 291% que el ingreso por habitante promedio de los hablantes seseantes. Los hablantes no yeístas, en cambio, tienen un ingreso por habitante promedio de 7.889 U\$S/año, y eso es un 44% menos que el ingreso por habitante promedio de los hablantes yeístas¹⁰.

Cuadro 8: PBI per cápita por característica (2008)

Característica / PBI pc	Sí	No
Seseo	11.649	33.921
Yeísmo	14.153	7.889
Aspriación de /s/	13.098	13.821
Rehilamiento de /3/	14.893	13.341
Aspriación de /x/	10.484	15.629

La relación entre ingreso per cápita e importancia relativa de las características fonéticas de los distintos dialectos del idioma español es algo que ya ha sido señalado por la literatura socio-lingüística. Lipski (1997), por ejemplo, ha sugerido que, en general, las características fonéticas de los países poderosos gozan de más prestigio que la pronunciación típica de los países más pobres, y que dicha conexión se extiende también a nivel de cada país entre las zonas más importantes y las más postergadas.

¹⁰ En este punto incide mucho el hecho de que la mayoría de los hablantes no yeístas corresponden al español andino tradicional y al español paraguayo, que son las dos variedades asociadas con un PBI per cápita menor.

Si tomamos las cifras del cuadro 8 y vemos qué características están asociadas con mayores niveles de PBI per cápita nos encontramos con el resultado esperable de que las mismas coinciden casi exactamente con las del hablante de español peninsular estándar, ya que los hablantes de mayor ingreso por habitante son los que poseen yeísmo pero no seseo, ni aspiración de /s/ ni aspiración de /x/¹¹.

Los datos puramente descriptivos que aparecen en los cuadros 4, 5, 6, 7 y 8, sin embargo, no nos dicen qué características son más importantes como señal de un ingreso medio más elevado, ni cómo dichas características interactúan entre sí. Una forma analítica de abordar dicho problema es a través de un análisis de regresión estadística, que tome como variable dependiente al PBI per cápita y como variables independientes a las distintas características fonéticas.

El análisis de regresión estadística es relativamente común en fonética para encontrar correlaciones entre variables de tipo acústico que tienen que ver con la caracterización de los sonidos¹². También es muy común en economía para explicar el comportamiento de variables como puede ser el PBI per cápita. No resulta en cambio habitual para correlacionar variables económicas con variables lingüísticas, si bien existen algunos trabajos que han intentado avanzar en dicha línea, sobre todo en lo que se refiere a la relación entre capacidades lingüísticas y nivel de ingreso, y a la relación entre similitudes lingüísticas y comercio internacional¹³. En este punto, sin embargo, es importante diferenciar entre modelos que intentan utilizar a las variables lingüísticas como variables explicativas de las variables económicas y modelos en los cuales lo que se intenta es simplemente correlacionar y medir el valor de determinada característica de la lengua como señal de una variable económica¹⁴. Esto último es lo que nos proponemos llevar a cabo en el ejercicio estadístico que se muestra a continuación, en el cual haremos una regresión que tiene la siguiente forma:

$$PBIPC = \alpha_0 + \alpha_1 * SESEO + \alpha_2 * YEISMO + \alpha_3 * ASPIRAC + \alpha_4 * REHIL + \alpha_5 * ASPIRX ;$$

donde PBIPC es la variable dependiente (ingreso per cápita); SESEO, YEISMO, ASPIRAC, REHIL y ASPIRX son variables dicotómicas o *dummy* que operan como variables independientes; y α_0 , α_1 , α_2 , α_3 , α_4 y α_5 son los coeficientes a estimar.

Como el lector probablemente imagine, las variables independientes de esta regresión hacen referencia a la presencia o ausencia de seseo, yeísmo, aspiración de /s/, rehilamiento de /ʒ/ y aspiración de /x/, y adoptan un valor igual a 1 cuando la característica está presente y un valor igual a 0 cuando está ausente. El total de observaciones utilizadas en la regresión es de 27, y corresponden a los 15 países en los cuales hemos identificado una única variedad de español (México, Guatemala, Honduras, El Salvador, Nicaragua, Costa Rica, Panamá, Cuba, República Dominicana, Puerto Rico, Venezuela, Ecuador, Chile, Uruguay y Paraguay) y a 12 observaciones más de los cinco países en los cuales hemos identificado más de una variedad (3 observaciones para España y Colombia, y 2 para Perú, Bolivia y la Argentina)¹⁵. La regresión se llevó a cabo utilizando dos métodos alternativos: uno de mínimos cuadrados ordinarios y otro de mínimos cuadrados en el cual el peso de las observaciones se ponderó por el número de habitantes asociado a cada

¹¹ La única excepción a esto se da respecto del rehilamiento del fonema /ʒ/, que es una característica que no aparece en el español peninsular norteño y está sin embargo asociada a hablantes que poseen un ingreso per cápita mayor que el promedio de hablantes que no rehilan. El rehilamiento del fonema /ʒ/, sin embargo, aparece en el español andaluz y en el español rioplatense, que son dos variedades cuyo ingreso per cápita también es mayor que el promedio.

¹² Véase, por ejemplo, Clopper y Pisoni (2004), que es un trabajo en el cual se utiliza el análisis de regresión para caracterizar las diferentes variedades de inglés norteamericano, o Parker (2008), en el cual se lo emplea para analizar la sonoridad de los fonemas del inglés, el español y el quichua.

¹³ Respecto del primero de dichos temas, el trabajo pionero es Chiswick y Miller (1995). En lo referido a la relación entre lengua y comercio internacional, el artículo más influyente parece ser el de Helliwell (1999). Para una buena reseña en español de esta literatura, véase Jiménez (2006).

¹⁴ En ese sentido, nuestro enfoque estará más ligado a la literatura económica sobre "precios hedónicos". Para una reseña de la misma, véase Nesheim (2006).

¹⁵ El listado completo de las observaciones aparece en el apéndice del presente trabajo.

observación. En ambos se aplicó una corrección por heterocedasticidad basada en la metodología de White (1980).

Los resultados de los análisis de regresión son los que aparecen en el cuadro 9. En él hemos expuesto los valores obtenidos para los coeficientes estimados, así como también el “valor-p” de dichos coeficientes, que nos muestra una estimación de la probabilidad de que los mismos sean en rigor iguales a cero. Cuando dicha probabilidad es muy baja, se dice que el coeficiente en cuestión es “significativamente distinto de cero”, y para eso se suele utilizar un umbral que va del 10% al 1%. También hemos reportado en el cuadro los valores del coeficiente de determinación “R cuadrado”, que mide la bondad de ajuste de la regresión (que es mayor cuanto más cerca de 1 se encuentre el coeficiente en cuestión).

Cuadro 9: Resultados de los análisis de regresión del PBI per cápita

Variable	Sin ponderador		Ponderado por población	
	Coeficiente	Valor-p	Coeficiente	Valor-p
Constante	28698.80	0.0000	29309.85	0.0000
Seseo	-27624.34	0.0000	-20173.17	0.0000
Yeísmo	5932.02	0.0038	5050.90	0.0200
Aspiración de /s/	2684.10	0.3085	494.38	0.6887
Rehilamiento de /ʒ/	4616.52	0.2474	876.23	0.6687
Aspiración de /x/	2999.09	0.2196	-3125.54	0.0195
R cuadrado	0.747006		0.981711	

Los resultados de los análisis llevados a cabo muestran que, como señal del ingreso por habitante de quien la posee, la única característica fonética que tiene un claro efecto negativo es el seseo, ya que su existencia reduce el valor esperado del PBI per cápita anual en una suma de U\$S 27.624 ó de U\$S 20.173 (según la regresión que utilizamos), y es siempre estadísticamente significativa al 1%. Controlando por la presencia de seseo, en cambio, otras características como el yeísmo, la aspiración de /s/ y el rehilamiento de /ʒ/ tienen un efecto positivo sobre el valor esperado del ingreso por habitante promedio.

Si comparamos los resultados de las dos regresiones, vemos además que los mismos son relativamente similares en términos cualitativos, pero que la bondad de ajuste de la regresión en la cual las observaciones han sido ponderadas por el número de habitantes es considerablemente mayor (puesto que su R cuadrado es de 0,9817, contra 0,7470 de la regresión sin ponderador). Tomando como mejores los números de la regresión ponderada, vemos además que las otras dos características relevantes, además de seseo, son el yeísmo, que tiene un coeficiente positivo de U\$S 5.051 anuales que es estadísticamente significativo al 2%, y la aspiración del fonema /x/, que tiene un coeficiente negativo de U\$S 3.126 anuales que también es estadísticamente significativo al 2%. Esto implica que, como señales de un nivel de ingreso más elevado, las tres características importantes serían la ausencia de seseo, la presencia de yeísmo y la ausencia de aspiración del fonema /x/, si bien la primera de ellas exhibe un coeficiente que es varias veces mayor en valor absoluto que los de las otras dos.

Los resultados que aparecen en el cuadro 9 nos reafirman la idea que mencionamos al comentar el cuadro 8, respecto de que las características fonéticas asociadas con un mayor nivel de ingreso coinciden en general con el español peninsular estándar. Según esta visión, lo más relevante como señal de un ingreso más elevado es distinguir entre los sonidos de los fonemas /s/ y /θ/, no distinguir entre los fonemas /ʒ/ y /ʎ/, y no aspirar el fonema /x/, y no serían en cambio de importancia significativa la aspiración del fonema /s/ ni el rehilamiento del fonema /ʒ/.

5. Conclusiones

De todo lo expuesto y analizado en el presente trabajo resulta posible extraer una serie de conclusiones generales sobre las características fonéticas que presentan las

distintas variedades regionales del idioma castellano, y de la importancia económica que dichas características tienen entre los hablantes de dicho idioma.

Una primera conclusión que puede extraerse es que, a efectos de describir las principales características y diferencias entre las regiones que hablan español, basta con identificar cinco características fonéticas (seseo, yeísmo, aspiración de /s/, rehilamiento de /ʒ/ y aspiración de /x/), cuya presencia o ausencia permite tipificar diez variedades regionales (castellano tradicional, peninsular estándar, andaluz-canario, mexicano-centroamericano, caribeño, andino tradicional, andino moderno, chileno, rioplatense y paraguayo) que difieren entre sí en al menos una de las características analizadas.

Utilizando datos que nos permiten aproximar la importancia demográfica y económica de las regiones en las cuales dichas variedades se hablan, resulta posible inferir la importancia de las características fonéticas descritas, llegándose a una tipificación del “hablante promedio” del idioma español (es decir, de aquél cuya habla contiene las características que son mayoritarias). Dicho hablante coincide exactamente con el español mexicano-centroamericano (es decir, con alguien que posee seseo y yeísmo, pero que no aspira el fonema /s/ ni el fonema /x/, ni rehíla el fonema /ʒ/).

Los datos de número de habitantes y de PBI sirven también para hallar el ingreso promedio per cápita de las distintas regiones en las que se habla español y, a través de ello, inferir un valor económico de las características fonéticas. Haciendo estimaciones del PBI per cápita promedio de los hablantes que presentan y que no presentan cada una de las cinco características fonéticas estudiadas, vemos que la que está asociada con un ingreso por habitante mayor es la ausencia de seseo, y que además es mayor el ingreso medio de los hablantes yeístas (respecto de los no yeístas), así como el de los hablantes que rehílan el fonema /ʒ/ (respecto de los que no lo hacen), y el de los hablantes que no aspiran el fonema /s/ ni el fonema /x/ (respecto de los que sí lo hacen).

Una manera más analítica de cuantificar la relación entre ingreso por habitante y características fonéticas es la que surge de una de regresión estadística entre el PBI per cápita y cinco variables *dummy* que representan a las características descritas, ponderado por la población asociada con cada observación. Del mismo surge que las tres características significativas como señales de un ingreso por habitante más alto son la ausencia de seseo, la presencia de yeísmo y la ausencia de aspiración del fonema /x/. La primera de ellas está asociada con un aumento del valor esperado del ingreso per cápita anual de más de U\$S 20.000, la segunda con un incremento de algo más de U\$S 5.000, y la tercera con un aumento de menos de U\$S 3.200. Las tres son características fonéticas que aparecen en el español peninsular estándar, siendo la más importante de ellas (la distinción entre los fonemas /s/ y /θ/) la que se encuentra más directamente asociada con el castellano que se habla en la zona del norte y centro de España.

Referencias bibliográficas

- Avila, Raúl (2003). “La pronunciación del español: medios de difusión masiva y norma culta”; *Nueva Revista de Filología Hispánica*, vol 51, pp 57-79.
- Caballero, Mónica; Moreno, Asunción y Nogueiras, Albino (2009). “Multidialectal Spanish Acousting Modeling for Speech Recognition”; *Speech Communication*, vol 51, pp 217-229.
- Chiswick, B. R. y Miller, P. W. (1995). “The Endogeneity Between Language and Earnings”; *Journal of Labour Economics*, vol 13, pp 31-57.
- Clopper, Cynthia y Pisoni, David (2004). “Some Acoustic Cues for the Perceptual Categorization of American English Regional Dialects”; *Journal of Phonetics*, vol 32, pp 111-140.
- Colantoni, Laura (2008). “Variación micro y macro fonética en español”; *Estudios de Fonética Experimental*, vol 17, pp 65-104.
- Coloma, Germán (2010). “Caracterización fonética de las variedades regionales del español y propuesta de transcripción simplificada”; mimeo, Universidad del CEMA.
- Helliwell, John (1999). “Language and Trade”; en Breton, A. (editor): *Exploring the*

- Economics of Language*. Toronto, University of Toronto.
- Jiménez, Juan Carlos (2006). "La economía de la lengua: una visión de conjunto", DT 01/06; Madrid, Instituto Complutense de Estudios Internacionales.
- Lipski, John (1997). "En busca de las normas fonéticas del español"; en Colombi, Cecilia y Alarcón, Francisco: *La enseñanza del español a hispanohablantes*, capítulo 8. Nueva York, Heath.
- Martínez, Eugenio y Fernández, Ana (2007). *Manual de fonética española*. Barcelona, Ariel.
- Moreno Fernández, Francisco y Otero, Jaime (2006). "Demografía de la lengua española", DT 03/06; Madrid, Instituto Complutense de Estudios Internacionales.
- Nesheim, Lars (2006). "Hedonic price functions", CEMMAP WP 18/06. Londres, University College.
- Parker, Steve (2008). "Sound Levels Protrusions as Physical Correlates of Sonority"; *Journal of Phonetics*, vol 36, pp 55-90.
- Piñeros, Carlos (2008). *Estructura de los sonidos del español*. Nueva York, Pearson.
- White, Halbert (1980). "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity"; *Econometrica*, vol 48, pp 817-838.

Fuentes de los datos utilizados

- Abeceb (2008). "Análisis del PBI per cápita de Argentina por provincias". Buenos Aires, Abeceb.com.
- Banco Mundial (2008). *Population and Gross Domestic Product 2008*. Washington DC, World Bank.
- CIA (2009). *The World Factbook*. Washington DC, Central Intelligence Agency.
- DANE (2007). *Cuentas departamentales*. Bogotá, Departamento Administrativo Nacional de Estadística.
- Eurostat (2001). *Regional Gross Domestic Product in the European Union*. Bruselas, Oficina de las Comunidades Europeas.
- INE (1990). *Producto bruto interno por departamento*. La Paz, Instituto Nacional de Estadística.
- INE (2008). *Proyecciones demográficas*. La Paz, Instituto Nacional de Estadística.
- INEI (2007). *Producto bruto interno por departamentos*. Lima, Instituto Nacional de Estadística e Informática.

Apéndice: Datos utilizados en las regresiones

PAIS (DIALECTO)	PBI	POBLAC	PBIPC	SESEO	YEISMO	ASPIRAC	REHIL	ASPIRX
España (CT)	83.659	2.888	28.966	0	0	0	0	0
España (PS)	1.111.790	32.354	34.363	0	1	0	0	0
España (AC)	247.488	10.314	23.995	1	1	1	1	1
Mexico	1.549.490	106.350	14.570	1	1	0	0	0
Guatemala	65.150	13.686	4.760	1	1	0	0	0
Nicaragua	15.240	5.667	2.689	1	1	0	0	0
El Salvador	41.707	6.134	6.799	1	1	0	0	0
Costa Rica	50.758	4.519	11.232	1	1	0	0	0
Honduras	28.776	7.319	3.932	1	1	0	0	0
Panamá	42.479	3.399	12.497	1	1	1	0	1
Cuba	110.800	11.205	9.888	1	1	1	0	1
Rep Dominicana	80.869	9.953	8.125	1	1	1	0	1
Puerto Rico	86.500	3.954	21.877	1	1	1	0	1
Venezuela	358.068	27.935	12.818	1	1	1	0	1
Colombia (AT)	106.706	12.354	8.637	1	0	0	0	1
Colombia (AM)	227.356	22.649	10.038	1	1	0	0	1
Colombia (CB)	61.924	10.009	6.187	1	1	1	0	1
Ecuador	108.044	13.481	8.015	1	1	0	0	1
Peru (AT)	44.298	9.176	4.828	1	0	0	0	1
Peru (AM)	201.060	19.661	10.226	1	1	0	0	1
Bolivia (AT)	20.425	4.880	4.185	1	0	0	0	1
Bolivia (AM)	21.037	4.814	4.370	1	1	0	0	1
Chile	242.583	16.804	14.436	1	1	1	0	0
Argentina (RP)	554.081	37.243	14.878	1	1	1	1	0
Argentina (PA)	16.765	2.640	6.349	1	0	1	1	0
Uruguay	42.489	3.334	12.744	1	1	1	1	0
Paraguay	29.345	6.238	4.704	1	0	1	1	0