

UNIVERSIDAD DEL CEMA
Buenos Aires
Argentina

Serie
DOCUMENTOS DE TRABAJO

Área: Lingüística

**LA FONOLOGÍA DEL
ESPAÑOL ARGENTINO**

Germán Coloma

Agosto 2022
Nro. 835

www.cema.edu.ar/publicaciones/doc_trabajo.html
UCEMA: Av. Córdoba 374, C1054AAP Buenos Aires, Argentina
ISSN 1668-4575 (impreso), ISSN 1668-4583 (en línea)
Editor: Jorge M. Streb; asistente editorial: Valeria Dowding <jae@cema.edu.ar>

La fonología del español argentino

Germán Coloma*

Resumen

Este trabajo describe la fonología del español hablado en la Argentina. Luego de considerar el inventario de fonemas de dicha variedad de español (tanto en lo referido a las vocales como a las consonantes) y sus principales alófonos y reglas fonológicas (incluyendo un apartado sobre la neutralización de fonemas y otro sobre el uso del acento como elemento distintivo), hay una descripción de las principales variaciones de pronunciación en las diferentes zonas geográficas de la Argentina. El trabajo incluye también una sección sobre la estructura silábica, que analiza las combinaciones permitidas de los distintos fonemas, así como también algunos aspectos ligados con procesos de resilabificación (cuando las palabras se pronuncian dentro de un discurso concatenado).

Palabras clave: fonología, español, Argentina, fonemas, variación, resilabificación.

The Phonology of Argentinian Spanish

Germán Coloma

Abstract

This paper describes the phonology of Spanish as it is spoken in Argentina. After considering the inventory of phonemes (both consonants and vowels) and their main allophones and phonological rules (including phoneme neutralizations and the use of stress as a distinctive element), there is a description of the most important variations of pronunciation in the different geographic regions of the country. The paper also includes a section about syllable structure, which deals with the admissible combinations of the different phonemes, and with some issues concerning resyllabification (when words are pronounced in the context of connected speech).

Keywords: phonology, Spanish, Argentina, phonemes, variation, resyllabification.

* Universidad del CEMA; Av. Córdoba 374, Buenos Aires, C1054AAP, Argentina. Correo electrónico: gcoloma@cema.edu.ar. El presente documento es básicamente una traducción de Coloma (2022). Agradezco los comentarios de Laura Colantoni y José Ignacio Hualde a dicho trabajo. Las opiniones expresadas en esta publicación son propias, y no representan necesariamente las de la Universidad del CEMA.

1. Introducción

El español o castellano es uno de los idiomas más hablados en el mundo. En la 25ª edición de *Ethnologue* (Eberhard, Simons & Fennig, 2022), por ejemplo, se estima que el número de hablantes nativos de español oscila alrededor de los 475 millones, y que el número total de hablantes es de unos 548 millones (una vez que le agregamos 73 millones de personas que hablan castellano como segunda lengua). El español, además, es un idioma “relativamente unificado”, en el sentido de que presenta un alto grado de inteligibilidad entre sus distintas variedades geográficas, con relativamente pocas variaciones en su morfología y en su sintaxis.

Las principales diferencias entre las variedades del castellano habladas en distintos lugares, sin embargo, tienen que ver con algunos aspectos del vocabulario y con la fonología del idioma. En este trabajo describiremos la fonología del español que se habla en la Argentina, país cuya población es de alrededor de 47 millones de personas (y más del 90% de ellas hablan castellano como primera lengua). Nuestro foco estará puesto en el habla de la zona de Buenos Aires, que es el principal centro político, económico y cultural del país, pero también incluiremos algunas observaciones acerca de variantes que se observan en otras partes de la Argentina, en especial en las regiones norte y oeste. La variedad principal descrita es asimismo muy cercana a la que se utiliza en la República Oriental del Uruguay, y por eso suele ser también denominada “español rioplatense”.

La estructura del presente documento es la siguiente. En la sección 2 describiremos el inventario de los fonemas del español argentino y sus principales rasgos distintivos. En la sección 3 analizaremos las principales reglas fonológicas que definen la pronunciación de dichos fonemas, sus posibles neutralizaciones, y las reglas relacionadas con la acentuación de las palabras. En la sección 4 nos concentraremos en las principales variaciones regionales de la fonología del castellano dentro de la Argentina, y en la sección 5 trataremos temas relativos a su estructura silábica (y a situaciones en las cuales dicha estructura se modifica como consecuencia de la concatenación de palabras dentro de un discurso hablado).

2. Inventario de fonemas

2.1. Consonantes

El inventario de las consonantes del español argentino estándar es el que aparece en el Cuadro 1. Consta de 17 fonemas que se articulan en cuatro lugares distintos (labial, dental, palatal y velar), utilizando siete modos de articulación diferentes. La mayoría de dichos fonemas, sin embargo, está representada por sonidos que son relativamente comunes en las distintas lenguas que se hablan en el mundo.

Cuadro 1. Fonemas consonánticos del español argentino

Modo/Lugar de articulación	Labial	Dental	Palatal	Velar
Oclusivo sordo	p	t	tʃ	k
Fricativo sordo	f	s	ʃ	x
Obstruyente sonoro	b	d		g
Nasal	m	n	ɲ	
Percusivo		r		
Vibrante		r		
Lateral		l		

El español argentino cuenta con una serie de cuatro consonantes oclusivas sordas (/p/, /t/, /tʃ/ y /k/) y cuatro fricativas sordas (/f/, /s/, /ʃ/ y /x/), y tiene también una tercera serie de “obstruyentes sonoras” (/b/, /d/ y /g/), que pueden pronunciarse de manera oclusiva ([b], [d], [g]) o fricativa ([β], [ð], [ɣ]). Estas últimas consonantes, inclusive, pueden también pronunciarse como “aproximantes”, es decir, como sonidos continuos sin fricción,¹ pero desde el punto de vista funcional o fonológico son siempre consonantes obstruyentes y no sonantes (véase Real Academia Española, 2011: pp 142-162).

Los fonemas sonantes del español argentino, por su parte, forman una serie de tres sonidos nasales (/m/, /n/ y /ɲ/) más otros tres sonidos “líquidos” (/r/, /r/ y /l/). El sonido nasal velar [ɲ] también existe en esta variedad del castellano, pero solo como un alófono de /n/ cuando dicho fonema aparece inmediatamente antes de otro fonema velar.

¹ Para una mayor explicación de este punto, véase Hualde & Colina, 2014: pp 50-53.

Las siguientes palabras-ejemplo, tomadas de Coloma (2018: pp 244), sirven para ilustrar la pronunciación de todos los fonemas consonánticos del español argentino:

p	pala	['pala]		
f	fama	['fama]		
b	bala	['bala]	β	pava
m	mapa	['mapa]		
t	tapa	['tapa]		
s	sala	['sala]	h	pasta
d	dama	['dama]	ð	nada
n	napa	['napa]	ŋ	manga
r	para	['para]		
r	parra	['para]		
l	lata	['lata]		
tʃ	chapa	['tʃapa]		
ʃ	llama	['ʃama]		
ɲ	ñata	['ɲata]		
k	cama	['kama]		
x	jarra	['xara]		
g	garra	['gara]	ɣ	maga

Aunque las distinciones fonológicas entre los puntos de articulación están limitadas a cuatro posiciones principales (que hemos definido en el Cuadro 1 como labial, dental, palatal y velar), algunos fonemas consonánticos tienen ciertas características articulatorias un poco más específicas. Por ejemplo, los fonemas /p/, /b/ y /m/ se articulan generalmente como sonidos bilabiales, en tanto que /f/ es típicamente labiodental. Del mismo modo, /t/, /d/ y /s/ son generalmente alveodentales, mientras que /n/, /l/, /r/ y /r/ son típicamente alveolares. Finalmente, /tʃ/ y /ʃ/ son posalveolares, en tanto que /ɲ/ es propiamente palatal. El fonema /tʃ/ se caracteriza también por tener una pronunciación “africada”, que implica el uso de un elemento oclusivo [t] seguido de un elemento fricativo [ʃ].

Las características mencionadas hasta aquí son comunes a casi todas las variedades de español que se hablan en el mundo, pero el castellano de la Argentina tiene también algunos rasgos que son típicos del español latinoamericano y diferentes del europeo. Una de ellas es la ausencia del fonema interdental fricativo sordo /θ/, que en el español rioplatense (y en el español latinoamericano, en general) se ha fusionado con el fonema /s/.

Este fenómeno suele ser designado en la literatura sobre lingüística hispánica con el nombre de “seseo” (véase, por ejemplo, Penny, 2004: pp 187-190, 220-222).

Otra característica del español de la Argentina, compartida con la mayor parte de los acentos latinoamericanos, es la pronunciación del sonido [s] como fricativo dorsal (en vez de fricativo apical, como ocurre en general en la península ibérica). El fonema /s/ tiene también un alófono glotal [h] en el español rioplatense, que se utiliza en algunos contextos específicos. Este rasgo es relativamente común en muchas variedades del castellano (por ejemplo, las que se hablan en Andalucía, Chile, el Caribe, etc.) y se lo conoce como “aspiración de /s/”. Por otro lado, en el español argentino, el sonido [h] nunca es un alófono del fonema /x/ (como sí ocurre en otras variedades tales como el español centroamericano, colombiano y caribeño).

Un último alófono importante que suele escucharse en el español argentino es el sonido posalveolar fricativo sonoro [ʒ], que se utiliza como una variante para el fonema /ʃ/.² De hecho, el sonido [ʒ] era la pronunciación más común de dicho fonema durante la primera mitad del siglo XX, y desde entonces ha venido siendo reemplazado gradualmente por el sonido [ʃ]. Esto es un fenómeno bastante particular del español rioplatense, ya que en otras variedades del castellano a este fonema se lo suele pronunciar como una consonante palatal fricativa sonora no asibilada [j], como una palatal sonora oclusiva [ɟ] o africada [jɟ], o inclusive como una semiconsonante “deslizante” [j].³ Hoy en día, la pronunciación más común en la zona de Buenos Aires parece ser [ʃ] (véase Rohena-Madrado, 2013: pp 47-55), si bien existe también una proporción importante de la población que utiliza el sonido [ʒ] en vez de [ʃ]. Esto es aún más evidente en otras regiones de la Argentina (por ejemplo, en la región patagónica), donde [ʒ] parece ser más común que [ʃ] (véase Fontanella, 2004: pp 214).

² En algunas descripciones relativamente antiguas del español bonaerense (p.ej., Fontanella, 1987: pp 144-150), /ʃ/ y /ʒ/ son analizados como fonemas distintos. En dichas descripciones, el fonema /ʒ/ se usa para la mayor parte de las palabras que se escriben con “y” o “ll” (p.ej., “yerno” [ˈʒerno], “lluvia” [ˈʒuβja]), en tanto que /ʃ/ es usado para algunos “préstamos léxicos” (p.ej., “shock” [ˈʃok], “champagne” [ʃamˈpan]).

³ Sobre este tema véase, por ejemplo, Clegg & Fails (2018: pp 302-310).

2.2. Vocales

Al igual que la mayor parte de las variedades del castellano, el español argentino solo tiene cinco fonemas vocálicos (ver Cuadro 2), que pueden aparecer tanto en sílabas acentuadas como no acentuadas. Las vocales posteriores (/u/ y /o/) suelen pronunciarse redondeando o “abocinando” los labios, en tanto que las otras tres vocales (/i/, /e/ y /a/) se pronuncian sin redondear los labios. No existe ninguna distinción relevante entre articulaciones más altas o más bajas de ninguno de estos fonemas, como sí ocurre en otras variedades de español (por ejemplo, en el sur de España).⁴

Cuadro 2. Fonemas vocálicos del español argentino

Modo/Lugar de articulación	Anterior	Central	Posterior
Cerrado	i		u
Medio	e		o
Abierto		a	

Las principales variantes que tienen las vocales en el español de la Argentina son la deslizante palatal [j] (para el fonema /i/) y la deslizante labiovelar [w] (para el fonema /u/). Estos sonidos pueden también escribirse como [i̟] y [u̟] (véase Hualde & Colina, 2014: pp 41-42). Algunos autores (p.ej., Hidalgo & Quilis, 2012: pp 162-171) usan los símbolos [j] y [w] cuando dichos sonidos aparecen antes del núcleo de una sílaba, y los símbolos [i̟] y [u̟] cuando aparecen después de dicho núcleo. Esto tiene que ver con la idea de que las deslizantes prenucleares son consideradas “semiconsonantes”, en tanto que las deslizantes posnucleares son consideradas “semivocales”.

Los fonemas /e/ y /o/ también admiten el uso de las deslizantes [ɛ̟] y [ɔ̟] como alófonos, en algunos diptongos que aparecen en el interior de las palabras y también en situaciones en las cuales se produce una “sinalefa” (es decir, cuando la última sílaba de una palabra y la primera sílaba de la siguiente palabra se pronuncian como si fueran una sola sílaba). Los siguientes ejemplos, tomados de Coloma (2018: pp 247) sirven para ilustrar la pronunciación de los fonemas vocálicos del español argentino:

⁴ Para un análisis de ese tema, véase Herrero (2019).

a	aso	['aso]			
e	eso	['eso]	ɛ	ese avión	['eseɔ 'βjon]
i	hizo	['iso]	j	hielo	['jelo]
o	oso	['oso]	ɔ̃	lago azul	['layɔa 'sul]
u	uso	['uso]	w	huevo	['weβo]

En castellano, las vocales adoptan una pronunciación parcialmente “nasalizada” si entran en contacto con alguna consonante nasal, pero no existe un contraste fonémico entre vocales orales y nasales (como sí ocurre en idiomas tales como el francés, el portugués y el guaraní). En particular, para el español bonaerense, se ha estimado que la nasalización es más probable si la vocal afectada es la /o/, si está rodeada de dos consonantes nasales, y si está localizada en la última sílaba de un enunciado (Bongiovanni, 2021: pp 5-6).

2.3. Rasgos distintivos

Los fonemas del español, tanto consonantes como vocales, pueden ser descriptos usando un conjunto de rasgos distintivos que permiten separar cada fonema de cualquier otro. En la gramática de la Real Academia Española (2011: pp 9-14) se utilizan 13 rasgos distintivos para caracterizar los fonemas de la lengua castellana. Algunos de ellos, sin embargo, son redundantes para llevar a cabo dicha caracterización, en particular para las variedades del español que no distinguen entre los fonemas /s/ y /θ/. En Coloma (2017: pp 70-74), por ejemplo, propusimos un esquema en el cual solo se necesitan 9 rasgos distintivos para caracterizar a las variedades latinoamericanas del castellano. Para el caso del español bonaerense, dichos rasgos son los que aparecen en el Cuadro 3.

Como puede verse, los rasgos distintivos que sirven para describir a los fonemas del español argentino son [+consonante], [+sonante], [+sonoro], [+alto], [+anterior], [+coronal], [+continuo], [+nasal] y [+lateral]. El primero de ellos sirve para distinguir entre los cinco fonemas vocálicos (que son todos [-consonantes]) de los fonemas consonánticos (que son todos [+consonantes]). Además, las vocales españolas son todas [+sonantes], [+sonoras] y [+continuas].

Cuadro 3. Rasgos distintivos de los fonemas del español argentino

Rasgo/Fonema	a	e	i	o	u	p	t	tʃ	k	f	s	ʃ	x	b	d	g	m	n	ɲ	ɾ	ʀ	l	
±consonante	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
±sonante	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+
±sonoro	+	+	+	+	+	-	-	-	-	-	-	±	-	+	+	+	+	+	+	+	+	+	+
±alto	-	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
±anterior	-	+	+	-	-	+	+	-	-	+	+	-	-	+	+	-	+	+	-	+	+	+	+
±coronal	+	+	+	-	-	-	+	+	-	-	+	+	-	-	+	-	-	+	+	+	+	+	+
±continuo	+	+	+	+	+	-	-	-	-	+	+	+	+	±	±	±	+	+	+	+	-	+	+
±nasal	±	±	±	±	±	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	-	-
±lateral	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+

Los tres rasgos que permiten diferenciar entre sí a los cinco fonemas vocálicos son [±alto], [±anterior] y [±coronal]. Mientras el fonema /a/ es [-alto], [-anterior] y [+coronal], /e/ es [-alto], [+anterior] y [+coronal]. Del mismo modo, el fonema /i/ es [+alto], [+anterior] y [+coronal], en tanto que /o/ es [-alto], [-anterior] y [-coronal]. Por último, el fonema /u/ es [+alto], [-anterior] y [-coronal].

Los rasgos [±anterior] y [±coronal] son también útiles para clasificar a las consonantes según su lugar de articulación. Así, mientras los fonemas labiales (/p/, /f/, /b/ y /m/) son todos [+anteriores] y [-coronales], los fonemas dentales (/t/, /s/, /d/, /n/, /ɾ/, /r/ y /l/) son [+anteriores] y [+coronales]. Del mismo modo, los fonemas palatales (/tʃ/, /ʃ/ y /ɲ/) son todos [-anteriores] y [+coronales], en tanto que los fonemas velares (/k/, /x/ y /g/) son [-anteriores] y [-coronales].

Los restantes rasgos distintivos nos permiten diferenciar entre consonantes que se articulan en los mismos lugares. Así, mientras las consonantes obstruyentes (/p/, /t/, /tʃ/, /k/, /f/, /s/, /ʃ/, /x/, /b/, /d/ y /g/) son todas [-sonantes], las consonantes sonantes (/m/, /n/, /ɲ/, /ɾ/, /r/ y /l/) son, por definición, [+sonantes]. Estos últimos fonemas son también [+sonoros], junto con las tres obstruyentes sonoras (/b/, /d/ y /g/). Por el contrario, las consonantes sordas /p/, /t/, /tʃ/, /k/, /f/, /s/ y /x/ son todas [-sonoras].

El rasgo [±continuo] es el que posibilita distinguir entre las oclusivas sordas /p/, /t/, /tʃ/ y /k/ (que son todas [-continuas]) y las fricativas sordas /f/, /s/, /ʃ/ y /x/ (que son todas

[+continuas]). Es también el que permite separar a la consonante percusiva /r/ (que es [-continua]) de la consonante vibrante /r/ (que es [+continua]).⁵ Esta última distinción es muy característica del idioma español (y relativamente rara en otras lenguas).

Nótese que el Cuadro 3 nos muestra también tres casos en los cuales el rasgo [±continuo] es irrelevante para describir un fonema consonántico. Esto ocurre con las obstruyentes sonoras /b/, /d/ y /g/, que pueden pronunciarse como [b], [d] y [g] (es decir, como [-continuas]) o como [β], [ð] y [ɣ] (es decir, como [+continuas]). Otra situación de irrelevancia para caracterizar un fonema consonántico en el español argentino ocurre cuando aplicamos el rasgo [±sonoro] al caso del fonema /ʃ/. Tal como hemos visto, dicho fonema es pronunciado como [ʃ] por algunos hablantes (es decir, [-sonoro]) y como [ʒ] por otros hablantes (es decir, [+sonoro]).

El último conjunto de rasgos que resulta útil para describir los fonemas consonánticos del español es el constituido por [±nasal] y [±lateral]. El primero de tales rasgos adopta el valor “+” para el caso de /m/, /n/ y /ɲ/, y el valor “-” para todas las otras consonantes. Los cinco fonemas vocálicos también pueden pronunciarse con cierto grado de nasalización, pero eso no es un rasgo relevante para caracterizar a las vocales del castellano (y por eso en las celdas que corresponden a esos fonemas hemos escrito el signo “±”). Finalmente, el rasgo [±lateral] es útil para describir a la consonante /l/ (que es [+lateral]), distinguiéndola de todos los otros fonemas (que son [-laterales]).

3. Reglas fonológicas

3.1. Fonemas consonánticos

Como hemos visto en la sección anterior, hay varios casos en los cuales los fonemas consonánticos del castellano se pronuncian utilizando distintos alófonos en diferentes situaciones. Esto permite establecer algunas reglas fonológicas acerca de la pronunciación esperada de dichos fonemas. En el español rioplatense, las reglas más importantes son las

⁵ La terminología utilizada para describir a los fonemas /r/ y /r/ (que emplea las palabras “percusiva” y “vibrante”) es la que usa la gramática de la Real Academia Española (2011: pp 245-252). Otros autores (véase, por ejemplo, Hidalgo & Quilis, 2012: pp 207-209) llaman “vibrante simple” a la consonante percusiva y “vibrante múltiple” a la consonante vibrante.

que tienen que ver con las consonantes obstruyentes sonoras (/b/, /d/ y /g/), con la consonante alveodental fricativa sorda (/s/) y con la consonante alveolar nasal (/n/).

Tal como hemos mencionado más arriba, las obstruyentes sonoras se pronuncian a veces como oclusivas ([b], [d], [g]) y a veces como fricativas o aproximantes ([β], [ð], [ɣ]). La articulación oclusiva se usa generalmente cuando tales fonemas aparecen después de una pausa (por ejemplo, al principio de una oración) o después de una consonante nasal. En el caso del fonema /d/, el alófono oclusivo se utiliza también cuando dicha consonante aparece después de /l/. La opción fricativa o aproximante, por su parte, se utiliza en los restantes casos. Ejemplos de todo esto son los que aparecen a continuación:

/b/	→	Buen día	[bwen 'dia]	→	[b] después de pausa
		envase	[em'base]	→	[b] después de nasal
		el vaso	[el 'βaso]	→	[β] después de lateral
		abajo	[a'βaxo]	→	[β] entre vocales
		los brazos	[loh 'βrasos]	→	[β] en otros casos
/d/	→	Dios mío	['djoh 'mio]	→	[d] después de pausa
		un deseo	[un de'seo]	→	[d] después de nasal
		el deseo	[el de'seo]	→	[d] después de lateral
		la desea	[la ðe'sea]	→	[ð] entre vocales
		las drogas	[lah 'ðroyas]	→	[ð] en otros casos
/g/	→	Gracias	['grasjas]	→	[g] después de pausa
		angosto	[aŋ'gohto]	→	[g] después de nasal
		el gasto	[el 'ɣahto]	→	[ɣ] después de lateral
		agosto	[a'ɣohto]	→	[ɣ] entre vocales
		los granos	[loh 'ɣranos]	→	[ɣ] en otros casos

Estas reglas, sin embargo, admiten algún grado de variación, en especial en lo referido al fonema /d/. Dicho fonema se omite a veces en el habla rápida o informal, y eso ocurre especialmente cuando el mismo aparece al final de una palabra (y, en algunos casos, entre vocales). Por ello es que en este caso puede enunciarse una “regla sociolingüística” adicional, que implica ejemplos como los siguientes:

/d/	→	la mitad	[la mi'tað]	→	[ð] en situación formal
		la mitad	[la mi'ta]	→	[Ø] en situación informal

Otro fonema que exhibe un buen grado de variación en el español de la Argentina es el fonema /s/, que en general se pronuncia como [h] cuando aparece antes de otra consonante (y después de una vocal), y como [s] en el resto de los casos. Ejemplos del uso de esta regla serían así los siguientes:

/s/	→	¡Salud!	[sa'luð]	→	[s] después de pausa
		paso	['paso]	→	[s] entre vocales
		pasto	['pahto]	→	[h] antes de consonante y después de vocal
		transporte	[trans'porte]	→	[s] entre consonantes
		perros	['peros]	→	[s] en otros casos

Esta regla de pronunciación del fonema /s/, sin embargo, también admite algunos cambios dependiendo del contexto. En el habla informal, por ejemplo, es común que las codas complejas, tales como la que aparece en la primera sílaba de la palabra “transporte”, se simplifiquen, y que en esa situación el fonema /s/ termine quedando antes de una consonante pero después de una vocal. Eso hace que dicho fonema también pueda pronunciarse como [h], como se muestra en el siguiente ejemplo:

/s/	→	transporte	[trans'porte]	→	[s] en situación formal
		transporte	[trah'porte]	→	[h] en situación informal

Adicionalmente, cuando el fonema /s/ aparece al final de una palabra, y la siguiente palabra comienza con una vocal, también es posible que su sonido se omita en el habla informal, con lo cual es posible encontrarse con ejemplos como los siguientes:

/s/	→	Vamos a ver	['bamosa 'βer]	→	[s] en situación formal
		Vamos a ver	['bamɔa 'βer]	→	[Ø] en situación informal ⁶

⁶ Una tercera alternativa (['bamoha 'βer]) también es posible aquí. La misma implica el uso de [h] en vez de [s] o [Ø]. Esta opción, sin embargo, no es tan común en el español de Buenos Aires, aunque sí lo es en el habla de otras regiones de la Argentina (véase, por ejemplo, Rojas, 2004, pp 170-171).

El fonema /n/, por último, también tiene un comportamiento particular que depende del contexto en el cual se ubica. Cuando el mismo aparece antes de una vocal o de una pausa, entonces siempre se pronuncia como [n]. Cuando aparece antes de una consonante, en cambio, adopta el punto de articulación de dicha consonante. Ejemplos de aplicación de esta regla son los siguientes:

/n/	→	en parte	[em 'parte]	→	[m] antes de consonante labial
		enjuague	[eŋ'xwaye]	→	[ŋ] antes de consonante velar
		entero	[en'tero]	→	[n] antes de otra consonante
		enano	[e'nano]	→	[n] entre vocales
		edición	[eði'sjon]	→	[n] en otros casos

3.2. Fonemas vocálicos

Tal como hemos mencionado en la sección 2, los fonemas vocálicos del castellano también pueden tener variación en su pronunciación. A excepción del fonema /a/ (cuyo único alófono relevante en el español rioplatense es la vocal central abierta [a]), todas las otras vocales pueden ser pronunciadas como “vocales plenas” ([i], [u], [e], [o]) o como deslizantes ([j], [w], [ɛ], [ɔ]). Para los casos de /i/ y /u/, el uso de uno u otro alófono es relativamente claro: las vocales plenas [i] y [u] deben usarse cuando el fonema es el núcleo de una sílaba, y las deslizantes [j] y [w] deben usarse en el resto de los casos. Esto es equivalente a decir que las deslizantes se usan en diptongos y triptongos en los cuales /i/ y /u/ no son las vocales nucleares, y que las vocales plenas se utilizan en todas las demás situaciones. Para comprender mejor esta regla, los siguientes ejemplos pueden ser de utilidad:

/i/	→	baile	['bajle]	→	[j] en diptongos decrecientes
		piano	['pjano]	→	[j] en diptongos crecientes
		fui mos	['fwimos]	→	[i] en diptongos donde /i/ es núcleo
		risa	['risa]	→	[i] en sílabas acentuadas
		mirar	[mi'rar]	→	[i] en otros casos

/u/	→	jaula	['xawla]	→	[w] en diptongos decrecientes
		cuatro	['kwatro]	→	[w] en diptongos crecientes
		viuda	['bjuða]	→	[u] en diptongos donde /u/ es núcleo
		ruta	['ruta]	→	[u] en sílabas acentuadas
		jugar	[xu'ɣar]	→	[u] en otros casos

Las vocales medias /e/ y /o/ también pueden pronunciarse como deslizantes en ciertas circunstancias.⁷ Esto ocurre, por ejemplo, si ambas vocales son adyacentes, y ninguna de ellas está acentuada. En ese caso, la primera vocal se pronuncia en general como una deslizante en el español argentino, tal como se observa aquí:

/e/	→	teología	[te̞olo'xia]	→	[e̞] en diptongos crecientes
/o/	→	coherencia	[koe'rensja]	→	[o̞] en diptongos crecientes

El mismo fenómeno ocurre si /e/ u /o/ aparecen inmediatamente antes de /a/, y ninguna de dichas vocales está acentuada, tal como ilustran los siguientes ejemplos:

/e/	→	realidad	[re̞ali'ðað]	→	[e̞] en diptongos crecientes
/o/	→	coacción	[kɔak'sjon]	→	[o̞] en diptongos crecientes

Por último, /e/ y /o/ también pueden pronunciarse como deslizantes cuando aparecen inmediatamente después de /a/, como ocurre en los siguientes casos:

/e/	→	aerosol	[a̞ero'sol]	→	[e̞] en diptongos decrecientes
/o/	→	ahogado	[a̞o'ɣaðo]	→	[o̞] en diptongos decrecientes

En todos los ejemplos con /e/ y /o/ señalados más arriba, sin embargo, el uso de alófonos deslizantes es opcional, ya que resulta perfectamente posible pronunciar esas palabras usando vocales plenas en vez de deslizantes (p.ej., [teolo'xia], [koe'rensja],

⁷ Para un análisis más completo de este fenómeno, véase Colantoni & Hualde (2016).

[reali'ðað], [koak'sjon], [aero'sol], [ao'ɣaðo]). Si ese es el caso, los correspondientes diptongos desaparecen y son reemplazados por hiatos (es decir, por construcciones en las cuales cada vocal está en una sílaba diferente). Las vocales plenas, además, son obligatorias en todas las demás situaciones. Por ejemplo:

/e/	→	cuerda	['kwerða]	→	[e] en diptongos donde /e/ es núcleo
		pena	['pena]	→	[e] en sílabas acentuadas
		penal	[pe'nal]	→	[e] en otros casos
/o/	→	androide	[an'drojðe]	→	[o] en diptongos donde /o/ es núcleo
		bota	['bota]	→	[o] en sílabas acentuadas
		votar	[bo'tar]	→	[o] en otros casos

3.3. Neutralización de contrastes fonémicos

La propiedad básica de un fonema para existir como una entidad independiente es su capacidad de contrastar con otros fonemas para expresar distintos significados lingüísticos. En español, por ejemplo, decimos que la consonante percusiva /r/ y la consonante vibrante /r/ son dos fonemas distintos porque podemos hallar numerosos “pares mínimos” (como “pera” [ˈpeɾa] y “perra” [ˈpeɾa]) en los cuales cambiar el sonido de [r] a [r] modifica el significado de una palabra. Los contrastes fonémicos, sin embargo, no necesariamente operan en todos los contextos posibles, y la oposición entre /r/ y /r/ en castellano es un buen ejemplo de esto, ya que solamente resulta relevante cuando dichos fonemas aparecen entre vocales dentro de una palabra. En todos los otros contextos, en cambio, el contraste entre /r/ y /r/ queda “neutralizado”, porque solamente uno de dichos fonemas puede aparecer, o porque cambiar un sonido por el otro no modifica el significado de la expresión lingüística que queremos producir.

En el caso de los fonemas /r/ y /r/, esta falta de contraste en la mayor parte de las posiciones posibles es aprovechada por la ortografía española oficial para prescribir que solo distingamos entre las representaciones de dichos fonemas cuando los mismos aparecen entre vocales dentro de una misma palabra, en cuyo caso debemos usar la letra “r” para representar al fonema /r/ y el dígrafo “rr” para representar al fonema /r/ (véase Real

Academia Española, 2012: pp 17-31). En todas las otras situaciones debemos en cambio emplear simplemente la letra “r”, puesto que no existe posibilidad de oposición entre /r/ y /r/, y es por lo tanto innecesario especificar cuál fonema debe utilizarse.⁸ Esta regla puede clarificarse mejor a través de los siguientes ejemplos:

Inicio de palabra	→	/r/	→	rojo	['roxo]
Inicio de sílaba después de consonante	→	/r/	→	enredo	[en'reðo]
Final de sílaba	→	/r/	→	arma	['arma]
Final de palabra	→	/r/	→	comer	[ko'mer]
Segundo elemento de inicio complejo	→	/r/	→	broma	['broma]
Inicio de sílaba después de vocal	→	/r/ o /r/	→	caro	['karo]
			→	carro	['karo]

Además de los fonemas /r/ y /r/, en español existen otras consonantes que están sujetas a procesos de neutralización de sus contrastes fonémicos.⁹ Un ejemplo de esto ocurre con las consonantes nasales /m/ y /n/. Estos fonemas contrastan entre sí cuando se encuentran al inicio de una sílaba, tanto al principio de una palabra como dentro de la misma. Esto puede verse en los siguientes ejemplos:

Inicio de palabra	→	/m/ o /n/	→	mudo	['muðo]
			→	nudo	['nuðo]
Inicio de sílaba dentro de una palabra	→	/m/ o /n/	→	rama	['rama]
			→	rana	['rana]

Sin embargo, cuando /m/ o /n/ están en posición de coda (es decir, al final de una sílaba) y la siguiente sílaba empieza con una consonante, su contraste se neutraliza, y su pronunciación pasa a estar regida por el punto de articulación del siguiente fonema. O sea que si la siguiente consonante es labial, tanto /m/ como /n/ se pronuncian [m]; si es velar, se pronuncian [ŋ]; y si es dental o palatal, se pronuncian [n]. Para entender mejor esto, los

⁸ Para una explicación más completa de esto, véase Torres-Tamarit (2020): pp 19-21.

⁹ Las vocales, en cambio, son siempre contrastantes en castellano en todas las posiciones en las cuales pueden aparecer, por lo cual no existen situaciones en las cuales dichos contrastes se neutralicen (al menos en el habla formal).

siguientes ejemplos pueden ser de utilidad:

Antes de consonante labial	→	[m]	→	álbum blanco	['alβum 'blaŋko]
			→	avión blanco	[a'βjom 'blaŋko]
Antes de consonante velar	→	[ŋ]	→	álbum grande	['alβuŋ 'grande]
			→	avión grande	[a'βjoŋ 'grande]
Antes de otras consonantes	→	[n]	→	álbum doble	['alβun 'doβle]
			→	avión doble	[a'βjon 'doβle]

La oposición entre /m/ y /n/, sin embargo, se restablece si dichos fonemas aparecen antes de una pausa, o antes de una palabra que comienza con una vocal. Por lo tanto, si bien estos fonemas están en coda, /m/ y /n/ contrastan entre sí en los siguientes ejemplos:

Antes de pausa	→	/m/ o /n/	→	Llegó el álbum.	[ʃe'yo el 'alβum]
			→	Llegó el avión.	[ʃe'yo el a'βjon]
Antes de vocal	→	/m/ o /n/	→	álbum azul	['alβum a'sul]
			→	avión azul	[a'βjon a'sul]

El español argentino se caracteriza también por poseer una neutralización adicional entre sonidos nasales, que es la que ocurre entre el fonema /ɲ/ y la combinación /ni/ (véase, por ejemplo, Bongiovanni, 2015). Esta neutralización no afecta a todos los hablantes del castellano rioplatense, pero es claramente perceptible en una fracción muy grande de la población cuando estos sonidos aparecen en el inicio de una sílaba y antes de las vocales /a/, /e/, /o/ o /u/. Un ejemplo de esto puede verse en el contraste entre las palabras “huraño” y “uranio”, que para muchas personas son homófonos (es decir, “suenan igual”) y pueden pronunciarse como [u'raɲo] o [u'ranjo]. Esto ha llevado a algunos autores (por ejemplo, Colantoni & Hualde, 2013) a proponer un análisis de la fonología del español argentino en el cual /ɲ/ ya no es más un fonema independiente. Sin embargo, la oposición entre /ɲ/ y /n/ todavía parece mantenerse cuando estos fonemas aparecen antes de /i/. Por ese hecho, hay palabras tales como “albañil” [alβa'ɲil] para las cuales no es posible utilizar [ɲj] en vez de [ɲ], en tanto que en otras palabras como “juvenil” [xuβe'nil] no resulta tampoco admisible

emplear el sonido [ɲ] en lugar de [n] o de [ni].

Un último grupo de neutralizaciones de contrastes fonémicos es el que ocurre entre las oclusivas sordas y las obstruyentes sonoras. Estas aparecen en los casos relativamente raros en los cuales tales sonidos están en coda en posición interna de palabra. En dichas circunstancias, la neutralización ocurre entre /p/ y /b/, entre /t/ y /d/, y entre /k/ y /g/.¹⁰ Ejemplos de esos procesos son los que afectan a las siguientes palabras:

/p/ y /b/	→	colapso	[ko'lapso] o [ko'laβso]
	→	absurdo	[aβ'surðo] o [ap'surðo]
/t/ y /d/	→	logaritmo	[loɣa'ritmo] o [loɣa'riðmo]
	→	admirar	[aðmi'rar] o [atmi'rar]
/k/ y /g/	→	anécdota	[a'nekðota] o [a'neyðota]
	→	amígdala	[a'miɣðala] o [a'mikðala]

3.4. Reglas de acentuación

En español, el acento es un elemento que resulta distintivo o contrastante, y puede aparecer en la última sílaba de una palabra (p.ej., “poder” [po·'ðer]), en la penúltima sílaba (p.ej., “poderoso” [po·ðe·'ro·so]) o en la antepenúltima (p.ej., “político” [po·'li·ti·ko]). En el primero de tales casos, se dice que la palabra es “aguda”, en el segundo caso se dice que es “grave”, y en el tercero se dice que es “esdrújula”.

Puede ocurrir también que una palabra no tenga acento, en cuyo caso se dice que es “átona”. Esto suele suceder con algunas partículas tales como preposiciones, conjunciones, artículos y pronombres, cuando se las utiliza dentro de un discurso concatenado. Por ejemplo, en la frase “el perro de la mujer” [el 'pero ðe la mu'xer], los sustantivos “perro” y “mujer” están acentuados, pero los artículos “el” y “la” y la preposición “de” son átonos.

En castellano, es relativamente común encontrar contrastes entre palabras que solo

¹⁰ La gramática de la Real Academia Española (2011: pp 153-155) menciona también casos de posibles neutralizaciones entre varias consonantes al mismo tiempo (p.ej., “absoluto”, que podría pronunciarse como [aβso'luto], [apso'luto], [atso'luto], [aðso'luto], [akso'luto], [aɣso'luto], [afso'luto] o [axso'luto]). La mayor parte de estas alternativas, sin embargo, son muy poco comunes en el español de la Argentina.

difieren en la posición de la sílaba acentuada. Por ejemplo, el verbo “revolver” [re·βol·'βer] es una palabra aguda (es decir, acentuada en la última sílaba) mientras que el sustantivo “revólver” [re·'βol·βer] es una palabra grave (cuya única diferencia con “revolver” es que está acentuada en la penúltima sílaba). Este tipo de oposición entre palabras acentuadas en distintas sílabas es sumamente productivo en español, en especial en la conjugación de los verbos (véase, por ejemplo, Roca, 2020). Se da así que muchos verbos tienen instancias en las cuales desplazar el acento de una sílaba a otra sirve para cambiar la persona y/o el tiempo verbal, tal cual puede observarse en los siguientes ejemplos:

canta ['kan·ta] (3ra persona presente indicativo)	vs.	cantá [kan·'ta] (2da persona imperativo)
cante ['kan·te] (1ra/3ra persona presente subjuntivo)	vs.	canté [kan·'te] (1ra persona pretérito perfecto simple)
canto ['kan·to] (1ra persona presente indicativo)	vs.	cantó [kan·'to] (3ra persona pretérito perfecto simple)

El idioma español posee también algunos “contrastes triples” entre palabras que tienen los mismos fonemas en el mismo orden, y solo difieren en la sílaba acentuada. Un ejemplo de esto es el caso de “término” [ter·mi·no] vs. “termino” [ter·'mi·no] vs. “terminó” [ter·mi·'no]. Estas oposiciones suelen ocurrir en casos de palabras que pertenecen al mismo campo semántico pero cumplen diferentes funciones dentro de un texto (por ejemplo, un sustantivo y dos formas distintas de un verbo relacionado).

Cuando existen dudas acerca de la posición de la sílaba acentuada en una palabra, la ortografía oficial del castellano (Real Academia Española, 2012: pp 36-58) prescribe el uso de una “tilde” o “acento ortográfico” (´). Dicho acento se escribe justo arriba de la vocal que opera como núcleo de la sílaba acentuada. La ortografía española, sin embargo, tiene algunas reglas que permiten evitar el uso excesivo de la tilde, y estas son esencialmente las siguientes:

a) Las palabras agudas llevan tilde si su última letra es “a”, “e”, “i”, “o”, “u”, “n” o “s”, y no la llevan en caso contrario (p.ej., “revés” [re·'βes] vs. “rever” [re·'βer]).

- b) Las palabras graves llevan tilde si su última letra no es “a”, “e”, “i”, “o”, “u”, “n” o “s”, y no la llevan en caso contrario (p.ej., “hábil” [‘a·βil] vs. “aves” [‘a·βes]).
- c) Las palabras esdrújulas llevan siempre tilde, sin importar la letra con la cual terminen (p.ej., “módulo” [‘mo·ðu·lo], “déficit” [‘de·fi·sit]).
- d) Los hiatos obligatorios llevan tilde cuando puede confundírseles con diptongos (p.ej., “países” [pa·‘i·ses] vs. “paisajes” [paj·‘sa·xes]).
- e) En general, los monosílabos no llevan tilde, con independencia de que sean palabras acentuadas o átonas (p.ej., “con pan” [kom ‘pan]).¹¹

Cuando una forma verbal incorpora “pronombres enclíticos”, resulta posible que el acento termine cayendo en una sílaba que aparezca antes de la antepenúltima dentro de la nueva palabra que se forma. A estas palabras se las suele denominar “sobresdrújulas”, y la ortografía oficial española prescribe que todas ellas deben llevar acento ortográfico en la vocal nuclear de la sílaba acentuada de la forma verbal original. Ejemplos de este tipo de palabra son “traíganmelos” [‘traj·γan·me·los], “moviéndoselo” [mo·‘βjen·do·se·lo], etc. Una regla general respecto de estas construcciones es que el acento nunca puede caer en un pronombre enclítico, y esa es básicamente la causa por la cual algunas de estas palabras se vuelven sobresdrújulas.

En el español argentino, sin embargo, existe la costumbre de acentuar a los pronombres enclíticos en algunas situaciones. Esto nunca sucede si la forma verbal aumentada por el pronombre enclítico se convierte en una palabra grave, pero sí ocurre muy frecuentemente cuando dicha palabra se vuelve esdrújula o sobresdrújula (lo cual es más probable en casos con dos pronombres enclíticos consecutivos). Algunos ejemplos de esto son los siguientes:

dejalo	[de·‘xa·lo]
dejámelo	[de·‘xa·me·lo] o [de·xa·me·‘lo]
déjenlo	[‘de·xen·lo] o [de·xen·‘lo]
déjenmelo	[‘de·xen·me·lo] o [de·xen·me·‘lo]

¹¹ Esta regla tiene algunas excepciones en casos en los cuales existen dos palabras con los mismos fonemas, pero una de ellas es tónica y la otra es átona (p.ej., “sí” [‘si] vs. “si” [si]). En dichos casos, la palabra tónica puede llevar tilde, y al mismo se lo conoce en la literatura lingüística hispánica como “acento diacrítico”.

Nótese aquí que para la palabra “dejalo”, que es grave, la única pronunciación admisible es la esperada por la ortografía oficial española (o sea, [de·'xa·lo]). Cuando la forma verbal se vuelve esdrújula (p.ej., “dejámelo”, “déjenlo”) o sobresdrújula (p.ej., “déjenmelo”), el español argentino tiene la alternativa de desplazar el acento hacia la última sílaba de la palabra, que se transforma en “aguda *de facto*”. Esta alternativa implica que el último pronombre enclítico (que en este caso es “lo”) termina quedando acentuado, y esto parece ser una tendencia dominante en el castellano hablado en la Argentina (véase, por ejemplo, Colantoni & Cuervo, 2013).

4. Variaciones regionales

En la descripción del inventario de fonemas y de las reglas fonológicas que hicimos en las secciones anteriores, nos hemos enfocado en la variedad de español que se habla en la ciudad de Buenos Aires. Esta variedad puede considerarse también como representativa de la pronunciación utilizada en toda el área sudeste de la Argentina, formada por las provincias de la región pampeana (Buenos Aires, Entre Ríos, Santa Fe, La Pampa y la zona sur de Córdoba) y de la región patagónica (Neuquén, Río Negro, Chubut, Santa Cruz y Tierra del Fuego).

El resto del país, en cambio, exhibe ciertas diferencias en la articulación de algunos fonemas, y puede dividírsele en tres áreas dialectales adicionales (ver Mapa 1), que son el noroeste, el noreste y el oeste (Coloma, 2018: pp 245-246). La primera de tales regiones incluye a las provincias de Jujuy, Salta, Santiago del Estero, Tucumán, Catamarca, La Rioja y la zona norte de Córdoba, en tanto que la segunda está formada por las provincias de Misiones, Corrientes, Chaco y Formosa (en la frontera con Paraguay). La región oeste, por último, abarca las provincias de San Juan, Mendoza y San Luis (cercanas a Chile).

La principal diferencia entre los acentos de las dos regiones norteñas (noreste y noroeste) y el de la zona de Buenos Aires es la asibilación y fricativización del fonema /r/ (Colantoni, 2008: pp 83-85). En dichas regiones, ese fonema se pronuncia en general como una consonante fricativa alveopalatal sonora [z], en vez de ser una consonante vibrante [r].

Esta fricativización, sin embargo, no se extiende a la pronunciación del fonema /r/, que en general se articula de manera percusiva [r] en todo el territorio argentino.

Mapa 1. Áreas dialectales del español argentino



La diferente articulación del fonema /r/ en el norte argentino, que puede tener que ver con el contacto con idiomas tales como el quechua y el guaraní durante varios siglos, no afecta en general las reglas enunciadas anteriormente respecto de la distinción y/o

neutralización entre /r/ y /r/. Puede por lo tanto observarse el uso del sonido fricativo [ʒ] al inicio de palabra y al inicio de sílaba después de una consonante (p.ej., “rojo” [ʒoxo], “enredo” [enʒeðo]), el uso del sonido percusivo [r] en codas y como segundo elemento de un inicio de sílaba complejo (p.ej., “arma” [ʼarma], “comer” [koʼmer], “broma” [ʼbroma]), y la oposición entre ambos sonidos en posición interna entre vocales (p.ej., “caro” [ʼkaro] vs. “carro” [ʼkazo]).

En la región noreste, además, hay mucha gente que utiliza el sonido lateral palatal [ʎ] como un fonema adicional, para aquellas palabras que se escriben con el dígrafo “ll”. Esta pronunciación tiene una larga tradición en el idioma español, pero su uso se ha ido reduciendo de manera sistemática durante el siglo XX en muchos lugares en los cuales era común al inicio de ese siglo (por ejemplo, en el norte de España y en el centro de Colombia). Hoy en día sigue estando muy difundida en Paraguay, y en algunas partes de Bolivia, Perú y Ecuador (véase Real Academia Española, 2011: pp 226–227).

En el noreste argentino, los hablantes que usan el fonema /ʎ/ pronuncian en general el fonema /ʃ/ con una articulación posalveolar africada sonora [dʒ], y distinguen entre palabras tales como “haya” [ʼadʒa] y “halla” [ʼaʎa]. Pero el fonema /ʎ/ también parece estar cambiando su pronunciación en la región noreste, variando desde un sonido lateral a uno de tipo palatal fricativo no asibilado [j] (véase Colantoni, 2008: pp 73-75). Así que mientras para un grupo importante de hablantes se mantiene la distinción entre los fonemas /ʃ/ y /ʎ/, dicha distinción puede manifestarse como un contraste entre [dʒ] y [j] (p.ej., “haya” [ʼadʒa] vs. “halla” [ʼaja]).

La pronunciación palatal fricativa no asibilada [j] está también muy generalizada en la región oeste de la Argentina, en la cual es virtualmente la forma estándar para el fonema /ʃ/. En dicha región no existe oposición entre /ʃ/ y /ʎ/, por lo cual palabras como “haya” y “halla” se pronuncian ambas [ʼaja]. Esta forma de pronunciar también es habitual en la región noroeste, si bien en dicha zona existe una variación considerable entre [j] y [ʒ] como realizaciones del fonema /ʃ/ (Rojas, 2004: pp 170-171). Por lo tanto, una palabra como “yuyo”, que en el sudeste de la Argentina se pronuncia [ʼjufo] o [ʼzufo], en el oeste se pronuncia [ʼjujo], y en el noroeste se pronuncia [ʼjujo] o [ʼzufo].

Otra variante que ha sido reportada como habitual en la región oeste de la Argentina (véase, por ejemplo, Gurlekian, Colantoni & Torres, 2001) es el uso del sonido palatal fricativo sordo [ç] como un alófono de /x/, cuando dicho fonema aparece antes de /e/ o /i/ (p.ej., “ají” [a'çi]). Este rasgo es también común en el español chileno (véase Real Academia Española, 2011: pp 194-195).

5. Estructura silábica

5.1. Sílabas permitidas

El Atlas Mundial de Estructuras Lingüísticas (Dryer & Haspelmath, 2013, capítulo 12) clasifica al español como una lengua con una estructura silábica “moderadamente compleja”. Esto implica que las sílabas pueden tener codas e inicios complejos, pero limitados a pocas combinaciones de sonidos. También implica que el número de fonemas que pueden “caber” dentro de una sílaba es relativamente pequeño, y que no están permitidas las sílabas con codas o con inicios excesivamente extensos, que sí se admiten en idiomas como el inglés (en el cual son posibles palabras monosilábicas tales como *screen* ['skri:n] “pantalla”, *worlds* ['wɜ:ldz] “mundos” o *sprints* ['sprɪnts] “corridas”).

De cualquier modo, la gramática de la Real Academia Española (2011: pp 293-298) menciona que el castellano tiene 23 tipos distintos de sílabas, que van desde un mínimo de un fonema hasta un máximo de cinco fonemas. El tipo más simple consta de una sola vocal plena (V), mientras que los tipos más complejos incluyen también consonantes (C) y deslizantes (D), en combinaciones tales como CCVCC (p.ej., “transmisión” [trans·mi·'sjon]), CCVDC (p.ej., “claustro” [klaus·tro]) y CCDVC (p.ej., “triángulo” [trjan·gu·lo]).¹²

A pesar de todas estas posibilidades, la idea de que el español tiene una estructura silábica relativamente simple tiene que ver más con un hecho estadístico que con una regla

¹² En español existen también algunos casos extremadamente raros de sílabas con seis fonemas, que ocurren en palabras aumentadas por un prefijo. Un ejemplo de ello es la primera sílaba de la palabra “preinscribir” [prejns·kri·'βir], cuya estructura es CCVDCC. En España, uno puede hallar también sílabas con seis fonemas en algunas formas verbales de la segunda persona del plural, tales como la última sílaba de “ampliáis” [am·pljajs], cuya estructura es CCDVDC.

de carácter absoluto. De hecho, la mayor parte de las sílabas que se usan en las palabras del castellano están formadas por una consonante seguida de una vocal (CV), y la posibilidad de que existan codas e inicios silábicos complejos está limitada por un conjunto de “restricciones fonotácticas” que reducen grandemente el número de sílabas permitidas (véase Colina, 2020). En rigor, se ha estimado que el 57,45% de las sílabas de los textos escritos en español tienen una estructura de tipo CV (Arias Rodríguez, 2016: pp 29), y que el número total de sílabas permitidas diferentes es igual a 11.642 (Coloma, 2017: pp 29). Esta es una cifra elevada si la comparamos con idiomas cuya estructura silábica es muy simple (por ejemplo, con el koiari, que se habla en Nueva Guinea, y que solo tiene 70 sílabas permitidas) pero es mucho menor que el número de sílabas que tienen lenguas como el vietnamita, que admite 52.007 sílabas diferentes (Shosted, 2006: pp 37-38).

En español, por ejemplo, los inicios silábicos con dos consonantes solo pueden producirse si la segunda de dichas consonantes es /r/ o /l/. Más aún, las combinaciones admisibles se limitan a 13 posibilidades, que son /pr/ (p.ej., “preso” [ˈpre·so]), /tr/ (p.ej., “trozo” [ˈtro·so]), /kr/ (p.ej., “crudo” [ˈkru·ðo]), /fr/ (p.ej., “frío” [ˈfri·o]), /br/ (p.ej., “bruja” [ˈbru·xa]), /dr/ (p.ej., “drama” [ˈdra·ma]), /gr/ (p.ej., “grito” [ˈgri·to]), /pl/ (p.ej., “pluma” [ˈplu·ma]), /tl/ (p.ej., “atlas” [ˈa·tlas]), /kl/ (p.ej., “clavo” [ˈkla·βo]), /fl/ (p.ej., “flecha” [ˈfle·tʃa]), /bl/ (p.ej., “blando” [ˈblan·do]) y /gl/ (p.ej., “globo” [ˈglo·βo]).

Del mismo modo, las codas con dos consonantes también están severamente restringidas en castellano, limitándose a casos en los cuales la segunda consonante es el fonema /s/. Las posibilidades son aquí las siguientes: /bs/ (p.ej., “abstemio” [aβs·ˈte·mjo]), /ds/ (p.ej., “adscripto” [aðs·ˈkrip·to]), /ps/ (p.ej., “fórceps” [ˈfor·sep̺s]), /ts/ (p.ej., “superávits” [su·pe·ˈra·βits]), /ks/ (p.ej., “externo” [eks·ˈter·no]), /gs/ (p.ej., “blogs” [ˈbloɣ̺s]), /fs/ (p.ej., “chefs” [ˈtʃef̺s]), /ns/ (p.ej., “instinto” [ins·ˈtin·to]), /rs/ (p.ej., “perspectiva” [pers·pek·ˈti·βa]) y /ls/ (p.ej., “solsticio” [sols·ˈti·sjo]). También está el caso de la palabra “istmo” [ˈist̺·mo], cuya primera sílaba tiene la combinación /st/ en coda. Dicha combinación aparecía también en algunos latinismos que comenzaban con el prefijo “post” (p.ej., “postdata” [post̺·ˈða·ta]). Dichas palabras, sin embargo, hoy en día se escriben y se pronuncian sin la letra “t” (p.ej., “posdata” [poh·ˈða·ta]), por lo cual la coda formada por

el dígrafo /st/ es prácticamente inexistente en el español moderno.¹³

Las codas formadas por una sola consonante también están sujetas a ciertas restricciones dentro de la estructura silábica del castellano. Al final de palabra, por ejemplo, es usual encontrar alguno de los siguientes cinco fonemas: /n/ (p.ej., “camión” [ka·'mjɔn]), /s/ (p.ej., “atrás” [a·'tras]), /d/ (p.ej., “salud” [sa·'luð]), /r/ (p.ej., “lugar” [lu·'ɣaɾ]) y /l/ (p.ej., “papel” [pa·'pel]). Menos frecuentes, pero posibles, son las palabras terminadas en estos otros fonemas: /b/ (p.ej., “club” ['kluβ]), /p/ (p.ej., “kétchup” ['ke·tʃup]), /f/ (p.ej., “rosbif” [roh·'βif]), /m/ (p.ej., “tótem” ['to·tem]), /t/ (p.ej., “robot” [ro·'βot]), /tʃ/ (p.ej., “sándwich” [saŋ·witʃ]), /g/ (p.ej., “esmog” [eh·'moɣ]), /k/ (p.ej., “coñac” [ko·'ɲak]) y /x/ (p.ej., “reloj” [re·'lox]). El fonema /ʃ/ también es usado en el español argentino en las codas de algunas palabras tomadas en préstamo de otros idiomas (p.ej., “garage” [ga·'raʃ]), pero ni /r/ ni /p/ pueden aparecer en coda.

En posición de inicio silábico, en cambio, los 17 fonemas consonánticos del español son posibles y contrastan entre sí. La única restricción aquí es la ya mencionada ausencia de /r/ en posición inicial de palabra. El idioma castellano, además, tiene una clara preferencia por los inicios silábicos frente a las codas. Debido a ello, todas las secuencias formadas por una consonante entre dos vocales plenas (VCV) se silabifican como V·CV y no como VC·V, tal como puede observarse en los siguientes ejemplos:

amo	→	[a·mo] y no *[am·o]
paró	→	[pa·'ro] y no *[par·'o]
deshago	→	[de·'sa·ɣo] y no *[des·'aɣ·o]

Por último, puede observarse que los cinco fonemas vocálicos del español (/a/, /e/, /i/, /o/ y /u/) pueden todos ellos tomar el papel de núcleo silábico (p.ej., “mar” [maɾ], “red” [reð], “fin” [fɪn], “sol” [sol], “luz” [lus]), y en ese caso se pronuncian como vocales plenas. Todas las vocales salvo la /a/ pueden aparecer también como deslizantes cuando hay diptongos o triptongos. En esos casos, las restricciones implícitas son las siguientes:

¹³ La combinación /st/, sin embargo, es común en casos en los que /s/ es el último fonema de una sílaba y /t/ es el primer fonema de la sílaba siguiente (p.ej., “estado” [eh·ta·ðo], “lastimar” [lah·ti·mar], etc.).

- a) Cuando dos o más vocales están en la misma sílaba, la más abierta se pronuncia como una vocal plena, y las más cerradas como deslizantes. La vocal plena funciona entonces como núcleo de la sílaba. Por ejemplo, en la primera sílaba de la palabra “peine” [ˈpej.ne], /e/ es la vocal más abierta, y por lo tanto se pronuncia como una vocal plena [e]. Por el contrario, /i/ es más cerrada que /e/, y entonces se pronuncia como una deslizante [j].
- b) Dos vocales adyacentes no pueden estar en la misma sílaba si la más cerrada está acentuada. Por ejemplo, en la palabra “día” [ˈdi.a], /i/, que es más cerrada que /a/, está acentuada. Ocurre por lo tanto que /i/ y /a/ deben estar en sílabas diferentes, así que en este caso no hay un diptongo sino un hiato. Esta regla no depende del orden de las vocales, ya que también se cumple cuando la vocal más abierta (p.ej., /a/) precede a la más cerrada (p.ej., /i/). Así que en una palabra como “caí” [ka.ˈi] también hay hiato y no diptongo, ya que la vocal /i/ está acentuada, y por ende /a/ e /i/ deben estar en sílabas diferentes.
- c) Si dos vocales tienen el mismo grado de apertura (por ejemplo /e/ y /o/), y las dos están en la misma sílaba, la primera se pronuncia como una deslizante y la segunda como una vocal plena, y esta última es el núcleo de la sílaba. Por ejemplo, en la palabra “leopardo” [le.o.ˈpa.r.ðo], /e/ se pronuncia generalmente como una deslizante [ɛ] en el español rioplatense, y /o/ se pronuncia como una vocal plena [o]. Del mismo modo, en la palabra “poesía” [po.e.ˈsi.a], /o/ se pronuncia generalmente como una deslizante [ɔ] y /e/ como una vocal plena [e]. Lo mismo ocurre con la secuencia /iu/, la cual, en diptongos, siempre se pronuncia [ju] y nunca [iw] (p.ej., “ciudad” [sju.ˈða.ð], “triunfo” [ˈtrju.m.fo]).
- d) La única excepción a esta última regla aparece en unos pocos casos referidos al diptongo /ui/. En dicha secuencia, ambas vocales son cerradas, por lo cual /u/ debería pronunciarse como una deslizante [w] e /i/ debería pronunciarse como una vocal plena [i]. Eso es en rigor lo que sucede en la mayoría de los casos en el español argentino (p.ej., “fui” [ˈfwi], “cuidar” [kwi.ˈðar], “distribuir” [dih.tri.ˈβwiɾ]), pero no en el caso del adverbio “muy” [muj], en el cual el diptongo /ui/ no se pronuncia [wi] sino [uj]. Dicha pronunciación aparece también en algunas formas de algunos verbos terminados en “uir”, tales como “huiremos” [uj.ˈre.mos], “destruirían” [deh.truj.ˈri.an], etc.

5.2. Resilabificación

Los principios que rigen la estructura de las sílabas en español sufren algunos cambios importantes si dichas sílabas se analizan en el contexto de un discurso concatenado. En efecto, cuando las palabras se pronuncian dentro de una frase o enunciado, algunos fonemas pueden desaparecer, en tanto que otros pueden desplazarse de una palabra a otra. Considérese, por ejemplo, la siguiente oración:

Ellos sirvieron una ensalada de espinaca.

/e·ʃos sir·'bie·ron 'u·na en·sa·'la·da de es·pi·'na·ka/

Tal como puede observarse, esta oración está constituida por 6 palabras (léxicas), que pueden dividirse en 16 sílabas distintas. La pronunciación de este enunciado en el español argentino estándar, sin embargo, es la siguiente:

[e·ʃo sir·'βje·ro 'nu·naɐ̃n sa·'la·ða ðeh·pi·'na·ka]

Esta manera de pronunciar la oración implica usar solamente 14 sílabas, que se agrupan a su vez en 5 palabras (fonológicas). Dichas palabras son todas diferentes de las palabras léxicas originales, y tres de ellas ([sir'βjero], ['nunaɐ̃n] y [ðehpi'naka]) carecen de significado si se las analiza aisladamente. Más aún, las otras dos palabras sí tienen un significado específico en español, pero el mismo es distinto del original: [eʃo] (“ello”) no significa lo mismo que “ellos”, y [sa'laða] (“salada”) tampoco significa lo mismo que “ensalada”. Así que una persona que entiende un poco de castellano, pero que no está familiarizada con sus reglas de resilabificación, puede pensar que esta oración es acerca de “algo salado”, en vez de referirse a “gente que sirvió una ensalada”.

Para un hablante nativo de español, en cambio, [eʃo sir'βjero 'nunaɐ̃n sa'laða ðehpi'naka] no puede significar otra cosa que “ellos sirvieron una ensalada de espinaca”, puesto que todos los cambios de estructura silábica que ocurren en la pronunciación de dicha oración son completamente predecibles partiendo de un conjunto de principios que

regulan la resilabificación entre palabras adyacentes. Tales principios son esencialmente los siguientes:

a) Cuando una palabra termina en consonante, y la siguiente palabra empieza con la misma consonante, uno de dichos sonidos puede omitirse. Por eso, “ellos sirvieron” puede transformarse en [eʃo sir'βjeron], lo cual implica que dos consonantes consecutivas idénticas se convierten en una sola [s].

b) Cuando una palabra termina en consonante, y la siguiente palabra empieza con una vocal, la consonante puede ser absorbida por la primera sílaba de la segunda palabra. De este modo, “sirvieron una” se convierte en [sir'βjero 'nuna], y el sonido [n] (que pertenece a la palabra léxica “sirvieron”) se pronuncia como si fuera el inicio de la primera sílaba de la palabra “una” (que de ese modo se transforma en ['nuna]).

c) Cuando una palabra termina en vocal, y la siguiente palabra empieza con la misma vocal, uno de dichos sonidos puede omitirse. Esto hace que “de espinaca” se convierta en [ðehpi'naka], lo cual implica que dos vocales consecutivas idénticas se transforman en una sola [e], y por lo tanto las sílabas /de/ y /es/ se pronuncian como si fueran una única sílaba ([ðeh]).

d) Cuando una palabra termina en vocal, y la siguiente palabra empieza con una vocal diferente, la última sílaba de la primera palabra también puede fusionarse con la primera sílaba de la segunda palabra. Esto permite que “una ensalada” se transforme en [unaɛn sa'laða], y que la última sílaba de la primera palabra (/na/) se integre con la primera sílaba de la segunda palabra (/en/). En este caso, la vocal más cerrada (que es /e/) pasa a pronunciarse como una deslizante [ɛ], y la vocal más abierta (o sea, /a/) pasa a ser el núcleo de la nueva sílaba que se forma ([naɛn]).

Nótese que las dos primeras reglas de resilabificación que hemos enunciado implican situaciones en las cuales el número total de sílabas permanece inalterado (y por eso /e·ʃo sir·'bie·ron 'u·na/ tiene la misma cantidad de sílabas que [e·ʃo sir·'βje·ro 'nu·na]). Las otras dos reglas, en cambio, implican la aparición de “fusiones silábicas”, ya que dos sílabas que eran originalmente distintas cuando pronunciábamos las palabras aisladamente se convirtieron en una única sílaba dentro de un discurso concatenado. Esto

fue precisamente lo que hizo que “/de/ + /es/” se transformara en [ðeh], y que “/na/ + /en/” se transformara en [naen].¹⁴

Aunque las fusiones silábicas también pueden darse dentro de una misma palabra (véase, por ejemplo, Hualde, 2020: pp 164-166), sus características más interesantes aparecen cuando ocurren entre palabras adyacentes. En tales casos, resulta de utilidad establecer cuáles son las condiciones que inducen la ocurrencia de una fusión, y cuáles otras perjudican (o impiden) dicha ocurrencia. Un factor importante parece ser la acentuación de las sílabas, tal como muestran los siguientes ejemplos:

Canto afuera	/'kan·to a'fue·ra/	→	['kan·t̞a 'fwe·ra]
Canto algo	/'kan·to 'al·go/	→	['kan 't̞al·yo] o ['kan·to 'al·yo]
Cantó afuera	/kan·'to a'fue·ra/	→	[kan·'t̞a 'fwe·ra] o [kan·'to a'fwe·ra]
Cantó algo	/kan·'to 'al·go/	→	[kan·'to 'al·yo]

Como se ve, si la última sílaba de la primera palabra y la primera sílaba de la segunda palabra son ambas átonas (p.ej., “canto afuera”), la fusión entre tales sílabas es prácticamente obligatoria (['kan·t̞a 'fwe·ra]). Por el contrario, si las dos sílabas están acentuadas (p.ej., “cantó algo”), la fusión es casi imposible, y las sílabas permanecen separadas ([kan·'to 'al·yo]). Cuando una de dichas sílabas está acentuada y la otra no (p.ej., “canto algo”, “cantó afuera”), la fusión puede darse o no, dependiendo de la decisión del hablante (y, probablemente, del significado pragmático de la expresión que dicho hablante está produciendo). Hualde (2020: pp 172) menciona un estudio estadístico sobre el español hablado en Estados Unidos (Alba, 2006) que encontró que la fusión silábica es más frecuente si la última sílaba de la primera palabra está acentuada y la primera sílaba de la segunda palabra no lo está (p.ej., “cantó afuera”) que en el caso en el cual la última sílaba de la primera palabra no está acentuada y la primera sílaba de la segunda palabra sí lo está (p.ej., “canto algo”). Esto también parece ser así en el español rioplatense (y, probablemente, en la mayor parte de las variedades de español que se hablan en el mundo).

Las fusiones silábicas entre palabras adyacentes son claramente una fuente muy

¹⁴ Para una explicación más completa de este fenómeno, véase Hualde y Colina (2014: pp 76-78).

importante de diptongos en el castellano hablado. Por ejemplo, en la transcripción fonética de la fábula “El viento norte y el sol” que aparece en Coloma (2018: pp 248) podemos contabilizar un total de 32 diptongos, pero solo 19 de ellos ocurren dentro de las palabras que forman el texto de dicha fábula. Los restantes 13 diptongos se generan por medio de sinalefas, es decir, son el resultado de fusiones silábicas entre palabras adyacentes.

Las sinalefas son también la fuente más común de triptongos. De hecho, cuando uno quiere mencionar ejemplos de triptongos en castellano, queda limitado a un grupo pequeño de palabras relativamente poco comunes, tales como “buey” [ˈbwej] y “hioides” [ˈjoj·ðes]. En cambio, si uno incluye ejemplos en los cuales los triptongos surgen de fusiones silábicas, la lista se vuelve prácticamente infinita: “antigua usina” [an·ti·ɣwaw ˈsi·na], “continuó igual” [kon·ti·ˈnwoj ˈɣwal], “limpia y barre” [ˈlim·pjaɪ ˈβa·re], “pinos y eucaliptos” [ˈpi·no sjeu·ka·ˈlip·tos], etc.

El proceso de fusión silábica permite además generar instancias en las cuales aparecen casos de cuatro y hasta cinco vocales coexistiendo en la misma sílaba. Esto implica crear “cuatriptongos” y “quintiptongos”, tales como los que se dan en los siguientes ejemplos:

Nadie airea	/ˈna·die ai·ˈre·a/	→	[ˈna·ðjeaj·ˈre·a]
Actuó audazmente	/ak·ˈtuo au·das·ˈmen·te/	→	[ak·ˈtʷoaw ðah·ˈmen·te]
Volvió a Europa	/bol·ˈbio a eu·ˈro·pa/	→	[bol·ˈβjoaew·ˈro·pa]

Tal como puede verse, los dos primeros ejemplos (que generan sílabas con cuatro vocales) se dan en situaciones en las cuales la primera palabra termina en un diptongo creciente ([je], [wo]) y la segunda palabra empieza con un diptongo decreciente ([aj], [aw]). Para conseguir una fusión silábica, además, necesitamos que al menos una de las dos sílabas que se fusionan sea originalmente átona. En el primer ejemplo, ambas sílabas lo son; en el segundo, la última sílaba de la primera palabra está acentuada pero la primera sílaba de la segunda palabra es átona.

Más difícil aún es conseguir una sola sílaba con cinco vocales, que es lo que ocurre en nuestro tercer ejemplo. En ese caso hay un diptongo creciente ([jo]) seguido de una

vocal abierta ([a]) más un diptongo decreciente ([ew]). Otra cosa que necesitamos es que como máximo una de las tres sílabas esté originalmente acentuada (en este caso, la que lo está es la última sílaba de la primera palabra), y que ni el diptongo creciente ni el decreciente contengan la vocal [a]. Esto permite que la secuencia de vocales quede conformada del siguiente modo: “cerrada-media-abierta-media-cerrada”. Esa es una de las poquísimas combinaciones de cinco vocales que pueden coexistir dentro de una misma sílaba en español.¹⁵

Una última posibilidad de resilabificación, que no involucra fusiones silábicas, ocurre cuando una palabra termina en una coda compleja y la siguiente palabra empieza con una vocal. Eso es una situación muy rara en castellano, básicamente porque solo hay unas pocas palabras que finalizan en codas con dos consonantes. Sin embargo, es un caso interesante para mencionar, porque en él tiene lugar un fenómeno particular: solo la última consonante de la primera palabra es absorbida por la segunda palabra, mientras que la primera consonante permanece en coda. Ejemplos de esto son los siguientes:

vals austríaco	/ˈba <u>ɫ</u> s aus·'tri·a·ko/	→	['ba <u>ɫ</u> sa <u>w</u> h·'tri·a·ko]
tórax estrecho	/'to·ra <u>k</u> s es·'tre·tʃo/	→	['to·ra <u>k</u> se <u>h</u> ·'tre·tʃo]
bíceps izquierdo	/'bi·se <u>p</u> s is·'kjer·do/	→	['bi·se <u>p</u> si <u>h</u> ·'kjer·ðo]

Nótese que estas resilabificaciones imitan la manera en la cual se comportan las combinaciones de consonantes dentro de una misma palabra. En efecto, cuando encontramos combinaciones tales como /ls/, /ks/ y /ps/ dentro de una palabra, y las mismas aparecen justo antes de una vocal, las reglas de silabificación del español prescriben que la primera consonante está en coda y que la segunda es el inicio de la siguiente sílaba. Ejemplos de esto pueden hallarse en palabras como “balsa” ['baɫ·sa], “acceso” [ak·se·so] y “asepsia” [a·'sep·sja].

¹⁵ La secuencia “cerrada-cerrada-abierta-media-cerrada” también es posible, tal como lo ilustra otro ejemplo similar: “Fui a Europa” /fui a eu·'ro·pa/, que puede pronunciarse ['fwjaew·'ro·pa].

Referencias bibliográficas

- Alba, Matthew (2006). Accounting for variability in the production of Spanish vowel sequences. En Sagarra, Nuria & Jacqueline Toribio, eds: *Proceedings of the 9th Hispanic Linguistics Symposium*, pp 273–285.
- Arias Rodríguez, Iván (2016). Cálculo de frecuencias de aparición de fonemas y alófonos en español actual utilizando un transcriptor automático. *Loquens* 3(1), e029.
- Bongiovanni, Silvina (2015). Neutralización del contraste entre /ɲ/ y /nj/ en el español de Buenos Aires: Un estudio de percepción. *Signo y Seña* 27, pp 11–46.
- Bongiovanni, Silvina (2021). Acoustic investigation of anticipatory vowel nasalization in a Caribbean and a non-Caribbean dialect of Spanish. *Linguistics Vanguard* 7(1), 20200008.
- Clegg, Halvor & Willis Fails (2018). *Manual de fonética y fonología españolas*. Nueva York, Routledge.
- Colantoni, Laura (2008). Variación micro y macro fonética en español. *Estudios de Fonética Experimental* 17, pp 65–104.
- Colantoni, Laura & José Ignacio Hualde (2013). Variación fonológica en el español de la Argentina. En Colantoni, Laura & Celeste Rodríguez Louro, eds: *Perspectivas teóricas y experimentales sobre el español de la Argentina*, pp 21–36. Madrid, Iberoamericana Vervuert.
- Colantoni, Laura & José Ignacio Hualde (2016). Conditions on front mid-vowel gliding in Spanish. En Nuñez-Cedeño, Rafael, ed: *The Syllable and Stress*, pp 3-27. Boston, Mouton de Gruyter.
- Colantoni, Laura & María Cristina Cuervo (2013). Clíticos acentuados. En Colantoni & Rodríguez Louro, *op. cit.*, pp 143-157.
- Colina, Sonia (2020). Phonotactic constraints on syllable structure. En Colina, Sonia & Fernando Martínez-Gil, eds: *The Routledge Handbook of Spanish Phonology*, pp 131-144. Oxford, Routledge.
- Coloma, Germán (2017). *La complejidad de los idiomas*. Oxford, Peter Lang.
- Coloma, Germán (2018). Argentine Spanish. *Journal of the International Phonetic Association* 48(2), pp 243–250.
- Coloma, Germán (2022). The phonology of Argentine Spanish. En Jakobsen, Marcus, ed: *Sociolinguistics: Past, Present and Future Perspectives*. Nueva York, Nova Science.
- Dryer, Matthew & Martin Haspelmath (2013). *The World Atlas of Language Structures*. Leipzig, Max Planck Institute for Evolutionary Anthropology.
- Eberhard, David, Gary Simons & Charles Fennig (2022). *Ethnologue: Languages of the World*, 25th edition. Dallas, SIL International.

- Fontanella, Beatriz (1987). *El español bonaerense*. Buenos Aires, Hachette.
- Fontanella, Beatriz (2004). El español de la Patagonia. En Fontanella, Beatriz, ed: *El español de la Argentina y sus variedades regionales*, pp 209-219. Bahía Blanca, Asociación Rivadavia.
- Gurlekian, Jorge, Laura Colantoni & Humberto Torres (2001). El alfabeto fonético SAMPA y el diseño de corpora fonéticamente balanceados. *Fonoaudiológica* 47(3), pp 58–69.
- Herrero, Alfredo (2019). Catorce vocales del andaluz oriental. *Nueva Revista de Filología Hispánica* 67(2), pp 411-446.
- Hidalgo, Antonio & Mercedes Quilis (2012). *La voz del lenguaje: fonética y fonología del español*. Valencia, Tirant.
- Hualde, José Ignacio (2020). Syllable merger. En Colina & Martínez-Gil, *op. cit.*, pp 162-180.
- Hualde, José Ignacio & Sonia Colina (2014). *Los sonidos del español*. Cambridge, Cambridge University Press.
- Penny, Ralph (2004). *Variación y cambio en español*. Madrid, Gredos.
- Real Academia Española (2011). *Nueva gramática de la lengua española: fonética y fonología*. Madrid, Espasa-Calpe.
- Real Academia Española (2012). *Ortografía básica de la lengua española*. Madrid, Espasa-Calpe.
- Roca, Iggy (2020). Spanish verb and non-verb stress. En Colina & Martínez-Gil, *op. cit.*, pp 181-221.
- Rohena-Madrado, Marcos. 2013. Variación y cambio de sonoridad de la fricativa postalveolar del español de Buenos Aires. En Colantoni & Rodríguez Louro, *op. cit.*, 37-58.
- Rojas, Elena (2004). El español en el noroeste. En Fontanella, *op. cit.*, pp 161-187.
- Shosted, Ryan (2006). Correlating complexity: a typological approach. *Linguistic Typology* 10, pp 1-40.
- Torres-Tamarit, Francesc (2020). Phonemic contrast and neutralization. En Colina & Martínez-Gil, *op. cit.*, pp 3-33.