

Problemas teóricos y experimentales con las vocales del español

Andrea Cecilia Menegotto
Sofía Romanelli

Doctora en Lingüística (UBA) y profesora de Inglés (UNMDP). Fue becaria doctoral y postdoctoral en CONICET. Actualmente es docente e investigadora en la Universidad Nacional de Mar del Plata, Argentina.
Contato: sofiroma82@hotmail.com

Doctora en Lingüística (Universidad de Buenos Aires). Investigadora independiente en CONICET y profesora titular en la Universidad Nacional de Mar del Plata, Argentina.
Contato: acmenegotto@gmail.com
Argentina

Recebido em: 31 de agosto de 2019
Aceito em: 26 de setembro de 2019

PALABRAS CLAVE:

español rioplatense; vocales;
monoptongos; diptongos;
cambio espectral vocálico
inherente

Resumen: En este trabajo hacemos una descripción de las vocales del español rioplatense. Las describimos fonéticamente por sus propiedades acústicas y articulatorias presentando varias cartas de formantes. Ejemplificamos también el problema teórico de la identificación de los rasgos distintivos que permiten caracterizarlas. Finalmente, hacemos hincapié en dos problemas que consideramos cruciales para la descripción y la explicación fonética y fonológica: el problema del inventario —que surge básicamente de los presupuestos teóricos con los que se trabaje— y el problema de la distinción entre monoptongos y diptongos, habida cuenta de que los datos experimentales más actuales ponen en duda que las vocales del español sean monoptongos estables, ya que presentan dinamismo espectral inherente.

KEYWORDS:

River
Plate Spanish; vowels;
monophthongs; diphthongs;
vowel inherent spectral
change

Abstract: This article describes River Plate Spanish vowels. We describe vowels in terms of their acoustic and articulatory properties, and provide different vowel plots. The theoretical problem of identifying the distinctive features that characterize vowels was also exemplified. Finally, we give special emphasis into two issues we consider crucial for the phonetic and phonological description as well as explanation of vowels. The problem of vowel inventory- resulting from the theoretical assumptions underlying our work, and the problem of the distinction between monophthongs and diphthongs, as recent experimental data question the formant stability of Spanish monophthongs, providing evidence that vowels show spectral dynamic properties.

LA DESCRIPCIÓN FONÉTICA DE LAS VOCALES DEL ESPAÑOL

Las vocales son sonidos periódicos complejos que se producen gracias a los cambios en el movimiento vibratorio que el aire espirado produce en las cuerdas vocales, consideradas la fuente principal del sonido vocálico. Como cualquier sonido periódico, pueden describirse por medio de las frecuencias de las ondas que lo componen: cada vocal es un conjunto de formantes F0, F1, F2, F3, etc., y cada formante es una onda que resuena armónica y simultáneamente con las demás. Las frecuencias se miden en Hertz.

Las vibraciones producidas por la fuente atraviesan las cavidades del aparato fonador –el tubo traqueo faríngeo, la boca y la nariz principalmente– donde son amplificadas o atenuadas según las propiedades vibratorias de cada cavidad. Los órganos resonadores, entonces, son como filtros que dejan pasar ciertas frecuencias y bloquean otras. Además, las mismas cavidades resonadoras se convierten en fuente de sonido cuando las cuerdas no vibran y el aire encuentra alguna restricción en las distintas cavidades (Martínez Celdrán, 1998, 19).

La correspondencia entre las propiedades acústicas y las articulatorias de F1 y F2 es un supuesto aceptado en la disciplina: la posición de la mandíbula inferior y la distancia entre la lengua y el paladar que permite distinguir vocales altas, medias o bajas se corresponde con el valor de F1, y la articulación anterior o posterior se corresponde con F2. Entonces, cuanto más alta es la vocal, más bajo es el F1 y cuanto más atrás se articule, más bajo es el F2.

Por otra parte, F0 y F3 marcarían los límites de la percepción vocálica: mientras F0 indica el tono más bajo de cada hablante, F3 estaría relacionado más directamente con la longitud de su tracto vocal completo y, por lo

tanto, con su tono máximo (Idsardi, 2007). La frecuencia fundamental o F0 varía para cada hablante según la constitución de su aparato fonador, pero siempre dentro de cierto rango posible: los hablantes adultos hombres tienen rangos de F0 de entre 85 a 180 Hz, y las mujeres entre 165 y 255 Hz. (Baken, 1987, 177). El resto de los formantes son armónicos, es decir, proporcionales en relación al valor de F0.

Una forma tradicional de representar simultáneamente las propiedades acústicas y articulatorias de las vocales es a través de una carta de formantes—propuesta atribuida a Daniel Jones según Morales-Front (2014, 31)— con el F2 en el eje x y el F1 en el eje y y aumentando los valores hacia el punto de intersección (x, y), como se muestra en la Figura 1.

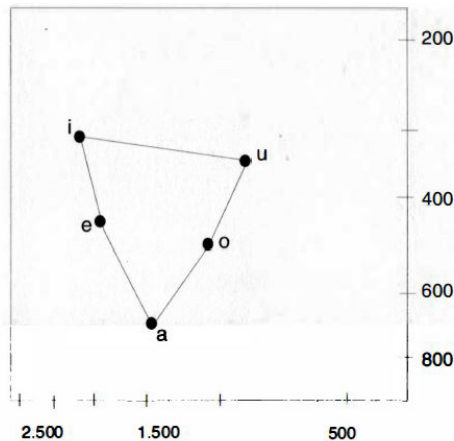


Figura 1: Forma tradicional de representar las vocales del español en una carta de formantes. Los valores del eje horizontal corresponden a F2, y el vertical a F1. Gráfico extraído de Martínez Celdrán (1998, 42).

Las cartas de formantes permiten representar más icónicamente la relación entre la articulación y las propiedades acústicas, al mostrar el triángulo vocálico. Metodológicamente, para construir las se suelen medir los formantes en el punto medio de la extensión de la vocal, para descartar los efectos de la coarticulación con el sonido que la precede o que le sigue. Así, las cartas producidas por diferentes autores, en distintas condiciones y con diversos recursos técnicos muestran resultados acústicamente muy similares: todas reproducen el triángulo vocálico. Cuando se representan los datos de múltiples hablantes, pueden encontrarse algunas zonas de superposición fonética entre pares de vocales, pero nunca entre tres. Esto es lo que se conoce como dispersión acústica de las vocales del español (Figura 2).

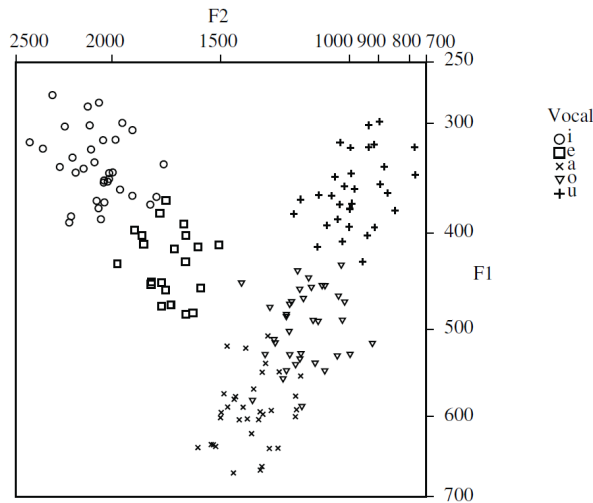


Figura 2: Dispersión de las vocales del español tomadas en 40 informantes hombres por medio de comunicación telefónica por celular. (Jiménez Gómez, 2011, 48).

Los datos de múltiples autores y su representación en carta de formantes son coincidentes en mostrar que las dos vocales altas del español, /i/ y /u/ tienen valores de F1 bastante cercanos, lo que se explica, precisamente, por ser altas, es decir, porque ambas se articulan con alguna parte de la lengua cerca del paladar. Se diferencian, sin embargo, por el valor de F2, que está vinculado a la anterioridad o posterioridad de la articulación lingual: cuanto más adelante está la lengua, más alta es la frecuencia de F2. Es decir que las vocales /i/ y /u/ se distinguen entre sí por los valores del segundo formante. Lo mismo sucede con la diferencia entre /e/ y /o/: ambas vocales tienen valores cercanos de F1 porque se articulan en la misma zona de altura media, pero tienen valores de F2 claramente diferentes pues una es anterior y la otra posterior: /e/ tiene un valor de F2 más alto que /o/. La /a/, en cambio, se distingue de todas las otras vocales tanto por sus valores de F1 como de F2, precisamente porque está más abajo y más en el centro que cualquiera de las otras: tiene el valor de F1 más alto de todas las vocales del español y un valor intermedio de F2.

En la segunda mitad del siglo XX, antes de que las computadoras personales fueran un recurso habitual para el trabajo del lingüista y del fonetista, se solía comparar los datos recogidos con el sistema de vocales cardinales. Así lo hicieron Guirao y Borzone (1975) cuando analizaron las propiedades acústicas de las vocales del español del Río de la Plata (Figura 3).

Con datos recogidos con los recursos técnicos de la actualidad (computadoras personales con buenos micrófonos y procesadores de audio y software de análisis del habla), muy diferentes de los utilizados por Guirao y Borzone en los setenta, los resultados de Romanelli et al (2018) coinciden en ubicar las

vocales /a/, /e/ y /o/ en las mismas regiones de la carta, aunque los autores incluyen, además, el contraste de vocales tónicas y átonas (Figura 4).

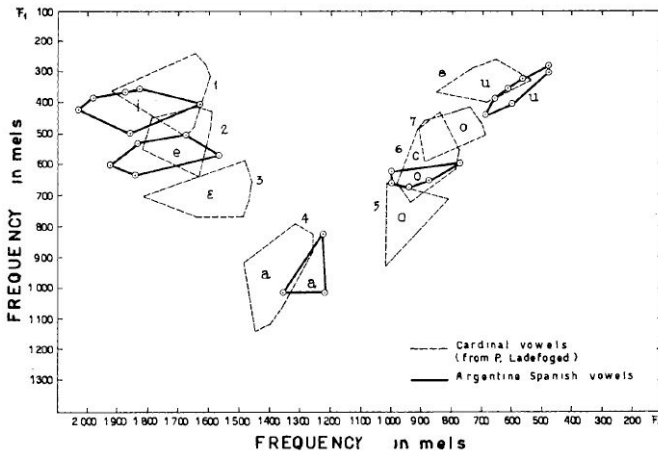


Figura 3: Carta de formantes de las vocales del español del Río de la Plata elaborada con datos de 5 hombres y 5 mujeres. Comparación con las vocales cardinales de Ladefoged. Guirao y Borzone (1975, 516).

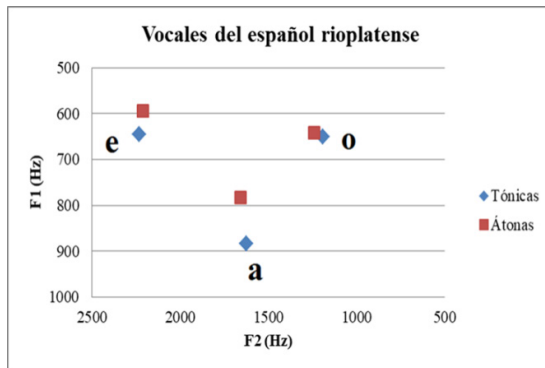


Figura 4: Vocales /a/ e /o/ tónicas y átonas del español rioplatense en una carta de formantes tradicional, elaborada con datos extraídos de Romanelli et al. (2018).

LA DESCRIPCIÓN FONOLÓGICA DE LAS VOCALES DEL ESPAÑOL

Desde los principios de fonología de Trubetzkoy en adelante, los rasgos forman parte de la descripción fonológica de cualquier lengua. La teoría de rasgos distintivos, desarrollada por Roman Jakobson y Moris Halle a partir del trabajo fundacional de Trubetzkoy para las lenguas eslavas, puede ser considerada como el principal aporte de la fonología a la teoría lingüística general. Jakobson y Halle (1955) marcaron una clara diferencia entre los rasgos distintivos como entidades fonológicas abstractas, y sus realizaciones articulatorias y acústicas concretas. Consideraron los rasgos fonológicos como constructos teóricos y a los rasgos fonéticos, en cambio, entidades concretas y medibles acústica o fisiológicamente.

Linguistic analysis gradually breaks down complex speech units into morphemes as the ultimate constituents endowed with proper meaning and dissolves these minutest semantic vehicles into their ultimate components, capable of differentiating morphemes from each other. These components are termed distinctive features. Correspondingly, two levels of language and linguistic analysis are to be kept apart: on the one hand, the semantic level involving both simple and complex meaningful units from the morpheme to the utterance and discourse and, on the other hand, the feature level concerned with simple and complex units which serve merely to differentiate, cement and partition or bring into relief the manifold meaningful units. (Jakobson & Halle, 1955, 12)

Los rasgos fonológicos cumplen dos funciones: distinguen sonidos contrastivamente y definen clases naturales. Casi todos los enfoques, desde el estructuralismo tradicional hasta las versiones más actuales de la teoría generativa y de la optimidad describen las unidades fonológicas por medio de rasgos y necesitan recurrir, de alguna manera, a la noción de sílaba.

Por ejemplo, siguiendo el esquema generativo clásico propuesto por Chomsky y Halle (1968), es posible caracterizar las vocales del español por el contraste de rasgos mayores primero, para separarlas de las deslizadas, las líquidas, las nasales y las obstruyentes por su sonoridad a partir de los rasgos [+ silábico], [+ consonántico] y [+ resonante]. Así, Prieto (2014) caracteriza las vocales como [+ silábicas], [-consonánticas] y [+ resonantes], distinguiéndolas fonológicamente de las deslizadas (las tradicionales semivocales o semiconsonantes i /u) por ser estas últimas [-silábicas], es decir, incapaces de funcionar como núcleo de sílaba.

El rasgo [+resonante] se asocia a la inexistencia de turbulencia en la columna de aire: en este sentido, tanto las vocales como las deslizadas, las líquidas y las nasales son [+resonantes]. El rasgo [+consonante] se emplea para distinguir vocoides (vocales y deslizantes) de consonantes. Las deslizadas (semivocales o semiconsonantes según la posición) son [-consonante] como las vocales. El rasgo [+ silábico] permite distinguir vocales de deslizadas, en tanto solo las primeras pueden ser núcleos vocálicos. Un diptongo, entonces, será una secuencia de deslizada y vocal o de vocal y deslizada.

Una vez identificadas como silábicas, resonantes y no consonantes, las vocales del español pueden subclasificarse fonológicamente con tres rasgos articulatorios: [+ alto], [+bajo] y [+retraído]

	I	E	A	O	U
Alto	+	-	-	-	+
Bajo	-	-	+	-	-
Retraído	-	-	-	+	+

Para definir [+ alto] y [+ bajo] es necesario asumir que hay un punto medio que corresponde a la posición neutral de la lengua: [+alto] indica que la lengua sube respecto de esa posición y [+bajo] que baja. Cuando la teoría parte del supuesto de rasgos binarios, [-alto] y [-bajo] se definen como instrucciones negativas: ni subir ni bajar la lengua, lo que resulta en una forma redundante de dar la misma instrucción de no mover la lengua.

Morales-Front (2014) y Prieto (2014) identifican el rasgo [+retraído] como el único rasgo necesario en función del punto de articulación vocálico. Si se considera que la posición neutral de la lengua no está ni adelantada ni retraída, entonces las vocales anteriores /e/, /i/ son [-retraídas], y las posteriores /o/, /u/ son [+retraídas].

También se ha propuesto el rasgo [+redondo]. Así, /o/ y /u/ se caracterizarían por tener el rasgo [+redondo] y /e/, /i/, /a/ por ser [-redondo]. Pero si se acepta que /a/ es [-retraído], es posible descartar la presencia de [redondo] entre los rasgos distintivos vocálicos: [+redondo] es predecible a partir de [+retraído], y [-redondo] a partir de [-retraído]: todas las vocales retraídas son redondas.

EL PROBLEMA DE LOS DIPTONGOS Y EL INVENTARIO VOCÁLICO

En síntesis, ya sea definiéndolas por sus propiedades fonéticas o fonológicas, está claro que las vocales son sonidos lingüísticos que se producen sin mayores obstrucciones en la salida del aire y que tienen la propiedad de ser el núcleo silábico.

Cuando definimos una vocal como anterior o baja estamos interpretando fonológicamente alguna propiedad articulatoria. Sin embargo, no queda demasiado claro qué entendemos acústica o articulatoriamente por sílaba; aunque creemos que sus bases fisiológicas están centradas en alguna propiedad vinculada con la actividad respiratoria, la importancia de la sílaba reside en que es la primera unidad fonológica propiamente dicha. Sobre ella se formulan las restricciones fonológicas más importantes de los sistemas lingüísticos:

Syllables are identifiable as the primary elements over which the rhythmic patterns of language can be observed, or the primary domain over which sequential constraints apply, or coarticulatory adjustments can be made. Vowels are defined by the physiological characteristic of their having no obstruction in the vocal tract, and by their function within a phonologically defined syllable.
Ladefoged y Maddieson 1996, 282)

Si las vocales son el pico de la sílaba, ¿qué pasa entonces con los diptongos? Una posibilidad es considerarlos una única vocal, es decir, un núcleo vocálico espectralmente dinámico. Así se describen, por ejemplo, varias de las vocales del inglés. Otra posibilidad es considerar que se trata de dos vocales diferentes (más precisamente, una deslizando y una vocal) unidas en la misma sílaba. Esto puede llevar a cambios drásticos en los inventarios vocálicos de las distintas lenguas: por ejemplo, Lee and Zee (2007) consideran que el chino estándar tiene 21 vocales, asumiendo que muchas de ellas son diptongos. Duanmu (2007), en cambio, considera que tiene tan solo 5.

En español es habitual considerar que un diptongo se compone de dos vocales diferentes unidas en la misma sílaba, siendo una de ellas una vocal

alta (i/u), aunque, más precisamente, no se trataría de una vocal alta sino de una deslizando, pues le faltaría el rasgo [+silábico]. De esta manera, el sistema vocálico del español tendría, además de las 5 vocales ya descriptas, 2 deslizando (semivocales) y 2 aproximantes (semiconsonantes).

La ventana de las tres sílabas es una de las propiedades fonológicas más características del español, y permite contrastar múltiples tripletes mínimos como *pálpito*, *palpito* y *palpitó*. Las vocales no solo son el núcleo silábico, sino que además llevan el acento, lo que permite que algunos consideren [+ acento] como un rasgo fonológico que reemplace a [+sílabo].

No sorprendentemente, la interacción de acento, deslizando y vocales plantea algunas paradojas en la formulación de reglas fonológicas del español. Como irónicamente menciona Harris sobre su propio trabajo, para lidiar con las reglas de acentuación en español es necesario dejar a los verbos o al resto de las categorías, fuera de las reglas:

“after brief background commentary, Harris (1983) announces that ‘verbs are excluded from discussion’ (p. 84). In Den Os and Kager (1986) verbs are not mentioned at all. Roca (1986) concentrates on postlexically- assigned non-primary stress; assignment of word stress in lexical categories falls outside its intended scope. Harris (1987) deals only with verbs and gives no clue as to how its analysis might be integrated with that of non-verbs in Harris (1983). Roca (1988) asserts starkly that “stress in verb paradigms obeys clearly distinct rules (...) and will not be dealt with here” (p. 398, fn. 6). On the other hand, Nuñez Cedeño (1985) and Otero (1986) take it as a desideratum to provide a unified formal description of stress in all morpho-syntactic categories, and they attempt to do so. However, the attempt does not succeed in these studies, which overlook certain fundamental and crucial data and/or present significantly flawed accounts of the range of data that they do deal with.” Harris, (1989, 241).

EL PROBLEMA DEL MOVIMIENTO DE FORMANTES EN MONOPTONGOS

Como ya señalamos, la metodología usual para la recolección de datos acústicos medía los formantes en un único punto correspondiente al núcleo vocálico, asumiendo que a lo largo de toda la duración del sonido vocálico nuclear las frecuencias serían invariables, y solo se registrarían modificaciones al inicio y al final por razones de coarticulación con los sonidos vecinos. Es decir, metodológicamente, asumíamos que las vocales del español eran monoptongos acústicamente estables, es decir que no tendrían ningún cambio espectral atribuible a las propiedades vocálicas *per se*. Cualquier cambio espectral era atribuible a los efectos de coarticulación.

Nearey y Assmann (1986) acuñaron el término *Vowel Inherent Spectral Change* (VISC) o Cambio Espectral Vocálico Inherente para referirse a las variaciones sistemáticas que se producen en los valores de los formantes de las vocales tensas y laxas del inglés a lo largo de su duración. El término describe, lo que es natural, los diptongos del inglés, pero también, sorprendentemente, los monoptongos. Para medir el VISC es necesario tomar los datos en, al menos, dos puntos. Hillenbrand, Getty, Clark & Wheeler (1995) midieron las frecuencias en diez puntos equidistantes de la duración en un amplio corpus de vocales de hablantes nativos de inglés (hombres, mujeres y niños), dando inicio a una muy rica generación de datos empíricos nuevos (Assman & Katz 2000; Elvin, Williams & Escudero 2016; Fox & Jacewicz 2009; Jacewicz, Fox & Salmons 2011a, 2011b; Morrison y Assmann 2012). De esta manera, se fue debilitando la certeza de que los monoptongos del inglés fueran sonidos estables sin cambio espectral, ya que el movimiento de formantes era evidente.

Fig. 1 Diphthongization of traditional monophthongs and phonetic diphthongs in Alberta English (vowels produced in isolation). *Arrow tails* Mean formant values measured at the beginning of the vowel. *Arrow heads* Mean formant values measured at the end of the vowel. Adapted from Nearey and Assmann (1986)

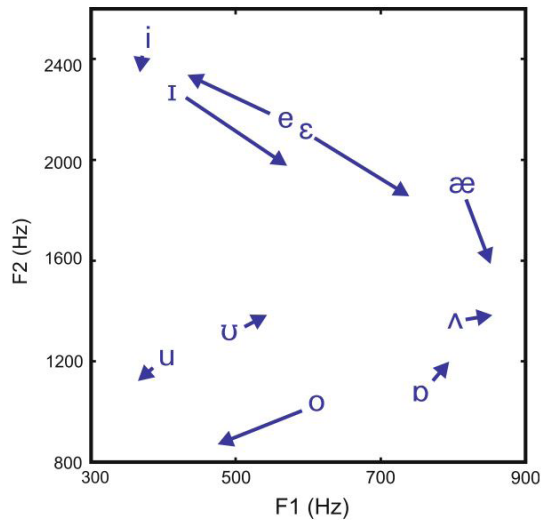


Figura 5: VISC de las vocales del inglés de Alberta reportadas por Morrison (2012): 33.

Obsérvese que en lugar de un solo punto por vocal, se representa un vector con origen y final, algunos más largos y otros notablemente breves¹

En el español rioplatense, Romanelli et al (2018) encontraron significativas muestras de VISC en las vocales /a/, /e/ y /o/ en posición final tomando valores formánticos en tres puntos: 25%, 50% y 75% de la duración de cada vocal. Representar el cambio espectral en cartas de formantes permitió ver que el movimiento tenía una cierta dirección: lo que antes se representaba como un único punto de intersección entre los valores de F1 y F2, ahora se

1 Obsérvese que el gráfico de Morrison no está configurado como una carta de formantes tradicional sino de una manera matemáticamente más transparente con los valores de F1 y F2 mínimos en el cruce de los ejes x e y, lo que permite calcular geométrica y algebraicamente la dirección de los movimientos. Para reinterpretar el gráfico como una carta de formantes es necesario girarlo y transponerlo.

representa como una línea que une tres puntos, cuya extensión representa, de alguna manera, la duración de la vocal y en la que se puede identificar con claridad hacia dónde se mueve cada formante, como se observa en la figura 6.

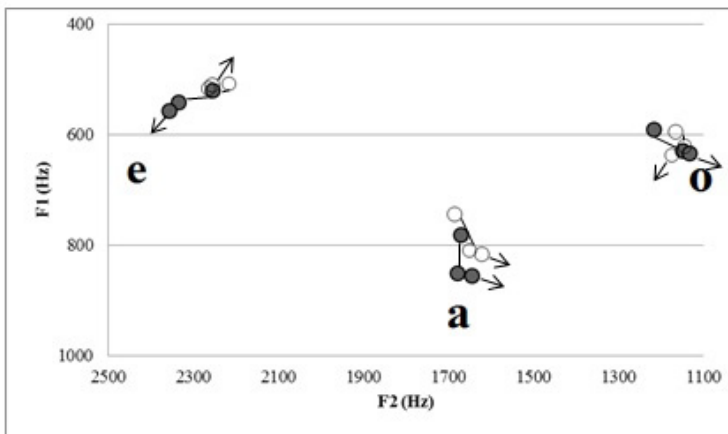


Figura 6: Carta de formantes de tres vocales tónicas (en gris) y átonas (en blanco) del español rioplatense producidas por diez hablantes nativos femeninos. (Romanelli y Menegotto ms).

Al confirmar con datos experimentales repetidos que las vocales /a/, /e/ y /o/ en posición final tónicas y átonas son dinámicas tanto en el F1 como en el F2, la caracterización de las vocales del español como monoptongos estables empieza a ser puesta en duda. Los resultados obtenidos hasta ahora muestran que, en estas tres vocales, sean tónicas o átonas, el movimiento espectral es mayor en la primera porción de la vocal en relación con la última.

Esto nos obliga a relativizar la caracterización tradicional de las vocales del español como acústicamente estables. Si bien comparativamente con

otras lenguas como el inglés parecen, efectivamente, ser menos dinámicas, existen cambios espectrales evidentes que las nuevas tecnologías disponibles nos permiten detectar y analizar.

CONCLUSIONES

En este trabajo describimos las vocales del español rioplatense por medio de sus propiedades fonéticas en cartas de formantes y ejemplificamos algunos de los problemas teóricos de la identificación de los rasgos distintivos que permiten caracterizarlas. Hicimos hincapié en dos asuntos que consideramos cruciales para la descripción y la explicación fonética y fonológica: el problema del inventario— que surge básicamente de los presupuestos teóricos con los que se trabaje— y el problema de la distinción entre monoptongos y diptongos, habida cuenta de que los datos experimentales más actuales ponen en duda que las vocales del español sean monoptongos estables. Nuestros datos muestran que los monoptongos /a/ /e/ y /o/, tónicos o átonos, tienen cambio espectral vocálico inherente.

Es necesario llevar adelante nuevas investigaciones en diferentes contextos y con múltiples condiciones diversas para determinar si los cambios espectrales que identificamos son generalizables a todos los contextos. Las cartas de formantes realizadas con datos de variados hablantes, en diferentes épocas y con diferentes recursos técnicos siempre mostraron que la medición de la vocal en el punto medio del núcleo silábico permitía reconstruir el triángulo vocálico de manera similar.

Es de esperar que las cartas de formantes realizadas con mediciones en tres o más puntos sean tan consistentes como las anteriores, y permitan avances teóricos significativos.

REFERENCIAS BIBLIOGRÁFICAS

- Assman, Peter; Katz, William. “Time-varying spectral change in vowels of children and adults”. In: *Journal of the Acoustical Society of America*, 108, 2000, 1856-1866.
- Baken, Ronald J. *Clinical Measurement of Speech and Voice*. London: Taylor and Francis Ltd., 1987.
- Chomsky, Noam; Halle, Morris. *The Sound Pattern of English*, Studies in Language, Longman, 1968.
- Duanmu, San. *The Phonology of Standard Chinese* (2nd edition). Oxford: Oxford University Press, 2007.
- Elvin, Jaydene; Williams, Daniel; Escudero, Paola. “Dynamic acoustic properties of monophthongs and diphthongs in Western Sydney Australian English”. In: *Journal of the Acoustical Society of America*, 140(1), 2016, 576-581.
- Fox, Robert A.; Jacewicz, Ewa. “Cross-dialectal variation in formant dynamics of American English vowels”. In: *Journal of the Acoustical Society of America*, 126, 2009, 2603-2618.
- Guirao, Miguelina; Borzone de Manrique, Ana M. “Identification of Argentine Spanish Vowels”. In: *Journal of Psycholinguistic Research*, 41975, 514-520
- Harris, James W. “How Different Is Verb Stress in Spanish?”. In: *Probus* 1 (3), 1989, 241-58.
- Hillenbrand, James; Getty, Laura A.; Clark, Michael J.; Wheeler, Kimberlee. “Acoustic characteristics of American English vowels”. In: *Journal of the Acoustical Society of America*, 97, 1995, 3099-3111.

- Idsardi, William J. "Some MEG Correlates for Distinctive Features." In: *Proceedings of 16th International Congress of Phonetic Sciences*, 2007.
- Jacewicz, Ewa; Fox, Robert A.; Salmons, Joseph. "Vowel change across three age groups of speakers in three regional varieties of American English". In: *Journal of Phonetics*, 39, 2011a, 683-693.
- Jacewicz, Ewa; Fox, Robert A.; Salmons, Joseph. "Cross-generational vowel change in American English". In: *Language Variation and Change*, 23, 2011b, 45-86.
- Jakobson, Roman; Halle, Morris. *Fundamentals of Language*, Mouton, 1955.
- Jiménez Gómez, Juan. "Estructura formántica y campo de dispersión de las vocales del español en la telefonía móvil". In: *Estudios fónicos I*, 2011, 39-58
- Ladefoged, Peter; Maddieson, Ian. *The Sounds of the World 's Languages*. Oxford: Blackwell Publishing Inc., 1996.
- Martínez Celdrán, Eugenio. *Análisis espectrográfico de los sonidos del habla*. Barcelona: Ariel, 1998.
- Morales-Front, Alfonso. "De la fonética descriptiva a los rasgos distintivos". In: Núñez-Cedeño, Rafael et al. (eds). *Fonología generativa contemporánea de la Lengua Española*. Washington, D.C.: Georgetown University Press, 2014, 25-43.
- Morrison, Geoffrey; Assmann, Peter. (eds). *Vowel inherent spectral change*. Springer, 2012
- Morrison, Geoffrey. "Theories of vowel inherent spectral change". In: Morrison y Assmann (eds). *Vowel Inherent Spectral Change*. Springer, 2012, 31-48.
- Nearey, Terrance M.; Assmann, Peter. "Modeling the role of inherent spectral change in vowel identification". In: *Journal of the Acoustical Society of America*, 80, 1986, 1297-1308.
- Prieto, Pilar. "La fonología". En Núñez-Cedeño, Rafael et al. (eds). *Fonología generativa contemporánea de la Lengua Española*, Washington, D.C.: Georgetown University Press, 2014, 1-24

Romanelli, Sofía; Menegotto, Andrea; Smyth, Ron. “Stress-Induced Acoustic Variation in L2 and L1 Spanish Vowels”. In: *Phonetica*, 75, 2018, 190-215

Romanelli, Sofía; Menegotto, Andrea. “Naturaleza dinámica de las vocales del español rioplatense”. In: *RASAL. Revista de la Sociedad Argentina de Estudios Lingüísticos*, 2019, 2: 82-98

Williams, Daniel; Escudero, Paola. “A cross-dialectal acoustic comparison of vowels in Northern and Southern British English”. In: *Journal of the Acoustical Society of America*, 136, 2014, 2751-2761.

Zee, Eric; Lee, Wai-Sum. “Vowel typology in Chinese”. In: *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrücken: Saarland University. 2007, 1429-1432