

## ALGUNOS DIPTONGOS DEL ESPAÑOL EN LA GRAN ÁREA METROPOLITANA DE COSTA RICA: ANÁLISIS FONÉTICO FONOLÓGICO

*On certain diphthongs in Costa Rica's great metropolitan area spanish:  
Phonetic and phonological analysis*

*Annette Calvo Shadid\**

*Manuel Ortega Rodríguez\*\**

### RESUMEN

Este artículo reporta la toma de datos y el análisis fonético articulatorio de la pronunciación de algunos diptongos del español de la Gran Área Metropolitana de Costa Rica. Hallamos que la consonantización en los diptongos en posición inicial no parece ser un rasgo característico de la población de esta región, al menos en el habla cuidadosa, y concluimos que fonológicamente existen diptongos en posición inicial en el español. Además, la consonantización, cuando ocurre, es más común ante *hue-* que ante *hie-*. Por otra parte, los hiatos en posición interna de palabra no se pronuncian como tales, sino que se suelen articular como diptongos, en tanto que el diptongo en *construir* se suele articular como hiato. Las excepciones importantes y la distribución por sexo, edad y escolaridad son también discutidas. De esta manera, esperamos aportar al denso debate sobre el problemático estatus fonológico en el español de diptongos en posición inicial, y de confluencia de vocales en general.

**Palabras clave:** español de Costa Rica, fonología, fonética, diptongos, acústica.

### ABSTRACT

This article reports on the data collection and articulatory phonetic analysis of the pronunciation of certain diphthongs in the Spanish of Costa Rica's Great Metropolitan Area. We found that consonantization of initial position diphthongs does not appear to be a typical feature of this region's population, at least in careful speech, and we conclude that there are, phonologically speaking, initial position diphthongs in Spanish. Moreover, consonantization, when it occurs, is more common in *hue-* than in *hie-*. On the other hand, middle position hiatuses are not pronounced as such, but they are usually articulated as diphthongs, while the diphthong in *construir* is usually articulated as a hiatus. Important exceptions, as well as sex, age and education-level distributions are also discussed. We hope thus to contribute to the dense debate on the controversial phonological status of Spanish initial position diphthongs and vowel confluence in general.

Key Words: Costa Rican Spanish, phonology, phonetics, diphthongs, acoustics.

---

\* Universidad de Costa Rica. Profesora catedrática. Decana Facultad de Letras. Costa Rica.  
Correo electrónico: acalsha@yahoo.com.mx

\*\* Universidad de Costa Rica. Profesor catedrático, Escuela de Física y Centro de Investigaciones Geofísicas.  
Correo electrónico: manuel.ortega@ucr.ac.cr  
Recepción: 20/3/16. Aceptación: 2/6/16.

## 1. Introducción

El estatus fonológico de los diptongos en posición inicial, particularmente, y de las confluencias de vocales, en general, ha sido motivo de debate y discusiones entre los diversos teóricos de la lengua española<sup>1</sup>. Calvo Shadid (2008) toca algunos puntos importantes sobre la fonética y la fonología de los segmentos considerados en español como semiconsonantes y semivocales, puesto que, según la investigadora, su estatus fonológico no está claro aún. Esto ocasiona una dificultad para explicar, tanto a nivel de modelo teórico como pedagógico, los fenómenos de consonantización de dichos segmentos en posición inicial de palabra, por ejemplo, o las diferencias entre hiatos y diptongos en la lengua en uso.

## 2. Marco teórico

### 2.1. Propuestas en torno a la fonética y la fonología de los diptongos

Haciendo un resumen de las propuestas más relevantes, se muestran aquellas que son pertinentes en esta investigación sobre el tema de la consonantización de los diptongos en posición inicial y de cuáles deslizadas se pueden dar en español<sup>2</sup>:

Navarro Tomás: (1918/61 y 1946/66) señala, sobre las características articulatorias de las semivocales y las semiconsonantes, que en las primeras su producción se realiza *con articulación momentánea*; de las segundas, indica que *la duración del sonido es brevísima*.

Al mencionar los diptongos en posición inicial de sílaba, señala que la /i/ y la /u/ se consonantizan en este contexto, y que en la pronunciación esmerada por influencia de la escritura se articulan como semiconsonantes.

Para Navarro Tomás hay, en la mayoría de los casos, una consonante en los diptongos que se encuentran en posición inicial de sílaba.

El autor considera, a partir del análisis que lleva a cabo, que los diptongos

conforman fonemas, y que las semivocales y semiconsonantes no se manifiestan en los vocablos con individualidad definida.

Por su parte, Alarcos (1961/83) concluye que los diptongos carecen de valor monofonemático y son combinaciones de los cinco fonemas vocálicos con otro elemento. Indica que en todos los casos observados, [j] e [i̯] son realizaciones de /i/, y que es /y/<sup>3</sup> toda realización que sea [ɔ̯], [y] e incluso [j].

Sobre la alternancia [i̯], [j] en palabras como rey [ˈre̯j], reyes [ˈre̯jes]; ay [ˈa̯j], ayes [ˈa̯jes], asigna estas variantes a /i/.

En algunas posiciones, considera el autor que /i/ y /j/ son conmutables, cuando aparecen entre los mismos fonemas, sobre todo en los casos de juntura (o sutura) de morfemas.

Respecto de [w] y [u̯], la situación se torna un tanto diferente que para [j] e [i̯], pues señala un refuerzo consonántico solo para [w], [gw], en posición inicial de palabra. Señala que tales variantes reforzadas [ɣw], [gw] se asemejan en su distribución a las variantes africadas (*fuertes*) de /j/, mientras [w] y [u̯] son paralelas a [j] e [i̯]; por eso, se deben considerar realizaciones de /u/. Propone una secuencia /gu/ en posición inicial de palabra (difonemática).

Cressey (1978:26) señala que tanto las semiconsonantes como las semivocales tienen una pronunciación idéntica, por lo que difiere de las afirmaciones de Navarro Tomás respecto de la diferencia de articulación entre [j] e [i̯]<sup>4</sup>.

Para Cressey las deslizadas<sup>5</sup> no forman núcleo de sílaba (son [-silábicas]), como sí lo hacen las vocales [+silábicas].

Respecto de las deslizadas que no son altas, Cressey va más allá de otros autores al señalar que el español presenta cinco deslizadas que corresponden a cada una de las cinco vocales. Así, habría cinco vocales plenas y cinco deslizadas correspondientes, como las clasifica en la figura que sigue:

Clasificación de todos los segmentos no consonánticos en el nivel fonético

[+silábico]	[-silábico]
i	j
u	w
e	ɛ
o	ɔ
a	a

Fuente: Calvo Shadid (2008), tomado de Cressey (1978:28) Figura 1.15.

Respecto de la consonantización de los diptongos en posición inicial de sílaba corresponde a una regla de consonantización que sigue a una regla de desvocalización de dichos segmentos antes o después de vocal.

Así, las derivaciones de *huevo* y *yerba* serían como se muestra en la siguiente figura:

	/uebo/	/ierba/
Regla (2.4)	w	j
Regla (2.5)	y <sup>w</sup>	y
	[y <sup>w</sup> e β o]	[y e r β a]

Fuente: Calvo Shadid (2008) de Cressey (1978:50)

La regla 2.4. es la siguiente:

(2.4) [-consonántico, +alto] ® [-silábico] / \_\_\_\_ [+ silábico] (a)  
 + silábico] \_\_\_\_\_ (b)

La regla 2.5 que consonantizaría en obstruyentes los segmentos [-consonánticos, -silábicos] es la siguiente:

(2.5) [-silábico, -consonántico] ® [+consonántico, -sonante]  
 / - \_\_\_\_\_

Según Cressey, las deslizadas fonéticas que derivan de vocales lo hacen por medio de una regla llamada *Low Level Allophonic Process Rule*. Algunas veces estas se pronuncian como deslizadas y otras como vocales; sin embargo, las deslizadas que se pronuncian siempre como tales, derivan de fonemas deslizados. Cressey propone entonces dos reglas de formación de deslizadas:

una ordenada antes de la regla de asignación del acento; la otra ordenada después de la regla de asignación del acento. Hay importantes diferencias respecto de ellas. La primera regla trata con los fonemas deslizados especificados por una *everywhere rule*. La siguiente regla, la *regla variable*, trata las deslizadas derivadas de vocales inacentuadas, como en los casos siguientes:

teatro 'theatre'	[te-a-tro],	[tja-tro]
poeta 'poet'	[po-e-ta],	[pwe-ta]
toalla 'towel'	[to-a-ʎa],	[twa-ʎa]
mi amor 'my love'	[mi-a-mor],	[mja-mor]
tu edad 'your age'	[tu-e-ðað],	[twe-ðað]
te adoro 'I adore you'	[te-a- ðo-ro],	[tja- ðo-ro]

Fuente: Calvo Shadid (2008), tomado de Cressey (1978:80)

La regla variable de Cressey establece que una vocal inacentuada se realizará como deslizada en contacto con otra vocal en el siguiente contexto: la operación de la regla se aplicará con mayor posibilidad a la más alta de las vocales, luego a las medias, luego a las bajas; la segunda parte indica que la regla se aplicará más probablemente donde no haya linde o linde morfológico entre los dos segmentos, y menos se aplicará si los dos segmentos están separados por linde de palabra. La aplicación de la regla se bloquea si el linde de frase fonético está presente entre dos vocales.

Luego, una regla que explicaría que una consonante laxa o aproximante se hace oclusiva después de pausa o después de una nasal o lateral homorgánica. Esto explicaría también el caso de un huevo [uŋ'gweβo]. Esta regla, denominada de *spirantización (spirantization of glides)* se ordena antes de la de formación de deslizadas.

Mosterín (1981) define el diptongo como un núcleo silábico formado por una semivocal [j] o [w], seguida de una vocal plena (en cuyo caso se habla de diptongo creciente) o por una vocal plena seguida por una semivocal [i] o [u] (en cuyo caso se habla de diptongo decreciente). Además de estos fonos, el autor distingue en

español otros bastante parecidos, pero que son claramente consonánticos y se pronuncian con cierta explosión o fricación del aire al pasar por una estrechez casi total de la boca formada por la lengua apoyada contra el paladar [j] o contra el velo [ɣw]. Estos fonos consonánticos (o semiconsonantes) se dan en posición inicial de sílaba: [jerba], *hierba*, [majó], *mayo*, [ɣweko], *hueco*, [aɣwekár], *ahuecar*. En Buenos Aires y en otros dialectos, al fono [j], corresponden frecuentemente el fono [ʒ] o [ʃ], según los hablantes.

D'Introno y otros (1995) reconocen la consonantización de los diptongos iniciales en español y hacen las siguientes propuestas:

Los fonemas que subyacen a las deslizadas son vocálicos. La condición para que una deslizada se consonantice, es que esta aparezca a inicio de sílaba, sin que esté precedida de consonante, pues no se consonantizaría. Esta condición es necesaria puesto que el Ataque es un constituyente constante de la sílaba, al igual que el núcleo vocálico, y a nivel fonológico contiene C. La sílaba por excelencia sería, entonces, constituida por CV. La deslizada es originalmente una vocal que *resilabea* cuando una sílaba tiene un ataque vacío, lo cual hace que dicha vocal se convierta en un ataque no vocálico.

Proponen que los lexemas de las palabras *hierro*, *hielo*, *hierba*, etc., alternan con *herramienta*, *hel-ado*, *herb-ívoros*, respectivamente, que tiene una vocal [e]. El lexema de estas palabras tiene a nivel fonológico una vocal /e/. Si la /e/ lleva acento, diptonga en [jé] debido a una regla de *Diptongación*, dando por ejemplo *hierro*. Después de esto, la deslizada, al encontrarse a comienzo de sílaba y al no estar seguida de límite de palabra, se convierte (obligatoriamente) en [j].

Por otra parte Núñez Cedeño y Morales-Front (1999), señalan que las deslizadas no forman grupo de ataque silábico con la consonante precedente, sino que son parte de la rima. Una excepción tiene que ver con las deslizadas en posición inicial de palabra, sin que preceda una consonante, que ocuparían el ataque de la sílaba, como la frecuente consonantización de [yélo], [wéso]. Otros casos de consonantización serían

respecto de la asimilación de nasales: [uñdzélo], [uñgwéso] y tras prefijos productivos como *des-*, como en *des.hie.lo*, que comparan con *de.sier.to*, y *re.sue.na* con *des.hue.sa*.

Respecto de su estatus fonológico, los autores señalan que en español hay un contraste fonológico entre vocales y deslizadas. Existe una distinción subyacente entre /i/, /u/, /j/, /w/, con la diferencia de que /i/, /u/ tienen una mora asociada, mientras que /j/, /w/ no la tienen. La sinalefa y la sinéresis se pueden conceptualizar como un proceso único de elisión de moras; es decir, una vocal pierde su mora cuando está en contacto con otra vocal.

## 2.2. Síntesis del apartado anterior

Sobre las propuestas de los autores, se consideran pertinentes para esta investigación las siguientes:

- a. Para los diptongos en posición inicial de sílaba, Navarro Tomás (1918) señala que la /i/ y la /u/ se consonantizan en este contexto en la mayoría de los casos, y que en la pronunciación esmerada por influencia de la escritura, se articulan como semiconsonantes.
- b. Alarcos (1961/83) señala un refuerzo consonántico solo para [w], [gw], en posición inicial de palabra. Señala que tales variantes reforzadas [ɣw], [gw] se asemejan en su distribución a las variantes africadas (*fuertes*) de /y/, mientras [w] y [ɥ] son paralelas a [j] e [i]; por eso, se deben considerar realizaciones de /u/. Propone una secuencia /qu/ en posición inicial de palabra (difonemática).
- c. Cressey va más allá de otros autores al señalar que el español presenta cinco deslizadas que corresponden a cada una de las cinco vocales. Así, habría cinco vocales plenas y cinco deslizadas correspondientes. Respecto de la consonantización de los diptongos en posición inicial de sílaba, para Cressey corresponde a una regla de consonantización que sigue a una regla

de desvocalización de dichos segmentos antes o después de vocal. La regla variable de Cressey establece que una vocal inacentuada se realizará como deslizada en contacto con otra vocal con mayor posibilidad a la más alta de las vocales, luego a las medias, luego a las bajas.

- d. Tanto Mosterín (1981) como D’Introno y otros (1995) reconocen la consonantización de los diptongos iniciales en español. D’Introno y otros (1995) plantean que los fonemas que subyacen a las deslizadas son vocálicos. La condición para que una deslizada se consonantice, es que esta aparezca a inicio de sílaba, sin que esté precedida de consonante, pues no se consonantizaría. Esta condición es necesaria puesto que el Ataque es un constituyente constante de la sílaba, al igual que el núcleo vocálico, y a nivel fonológico contiene C. La sílaba por excelencia sería, entonces, constituida por CV. La deslizada es originalmente una vocal que *resilabea* cuando una sílaba tiene un ataque vacío, lo cual hace que dicha vocal se convierta en un ataque no vocálico.
- e. Núñez Cedeño y Morales-Front (1999), señalan que las deslizadas no forman grupo de ataque silábico con la consonante precedente, sino que son parte de la rima. Una excepción tiene que ver con las deslizadas en posición inicial de palabra, sin que preceda una consonante, que ocuparían el ataque de la sílaba, como la frecuente consonantización de [yélo], [wéso]. Otros casos de consonantización serían respecto de la asimilación de nasales: [uńdʒélo], [uńgwéso] y tras prefijos productivos como *des-*, como en *des.hie.lo*, que comparan con *de.sier.to*, y *re.sue.na* con *des.hue.sa*. Los autores señalan que la sinalefa y la sinéresis se pueden conceptualizar como un proceso único de elisión de moras; es decir, una vocal pierde su mora cuando está en contacto con otra vocal.

### 2.3. Posición teórica adoptada en este estudio

Se tratará de establecer, con base en la investigación anterior (Calvo 2008), y con los nuevos datos recogidos, los siguientes supuestos teóricos:

1. Los segmentos [w] y [j] no forman núcleo silábico con la vocal del diptongo, puesto que las deslizadas en posición inicial de palabra pueden ocupar el ataque de la sílaba, tal y como se observa en su consonantización como en *huevo* o en *hierba*: [ˈgweβo~>ɣweβo, ⟨dʒerba~>jerba].
2. Las consonantizaciones en los diptongos en español no son tan frecuentes como lo proponía Navarro Tomás (/i/ y /u/ se consonantizan en posición inicial) y otros autores. En el 2008, la autora encontró poca cantidad de consonantizaciones. Se encontraron, en su gran mayoría, semiconsonantes (sonantes).
3. Al igual que lo señalan D’Introno y otros (1995), se supone que los diptongos en posición inicial de palabra presentan fonológicamente una vocal, a saber, /i/, /u/ que se desliza por contacto con la vocal plena y, en un proceso de resilabificación, la deslizada se puede convertir en una consonante prenuclear (en pocos casos, según la presente investigación). Se concuerda con Navarro Tomás (1946) en que en la pronunciación *esmerada* (cuidadosa), por influencia de la escritura, se articulan como semiconsonantes.
4. Los hiatos suelen asociarse en una sola mora y producirse como diptongos, una de cuyas vocales (media) se desliza en posición inacentuada y produce ascenso vocálico en algunos casos.
5. Las deslizadas se producen a partir de los fonemas vocálicos altos /i, u/, como lo proponen algunos de los autores arriba mencionados. El hecho de que los diptongos puedan desasociarse en ocasiones corrobora esta hipótesis; por

ejemplo, cuando se realiza un linde silábico entre ellos por asignación del acento: [ˈdja.βlo] frente a [di.ˈa.βlo] ‘diablo’, o entre [kons.tru.ˈir] y [kons.ˈtrwir] ‘construir’. Se concuerda con Núñez Cedeño y Morales Front (1999) en que la sinalefa y la sinéresis se pueden conceptualizar como un proceso de elisión de moras; una vocal puede perder su mora cuando está en contacto con otra vocal.

### 3. Metodología

#### 3.1. Las variables

En esta investigación no interesa llegar a datos numéricos contundentes, pues aunque se realizan mediciones y se extraen porcentajes, se refiere más a las cualidades del fenómeno de los diptongos y hiatos, y a un diagnóstico de la situación, que a una investigación cuantitativa y concluyente en cuanto a sus resultados. Sin embargo, se contemplaron las variables mínimas de sexo, dos grupos de edad, 19 a 35 y 36 y más, y dos niveles de instrucción, para clasificar preliminarmente la información en los grupos determinados. Se observará si estas variables mínimas, de una manera u otra, podrían afectar los resultados vinculados con la pronunciación de los diptongos. Además, debido a que Navarro Tomás habla de “pronunciación esmerada por influencia de la escritura”, se consideró dividir la muestra por nivel de instrucción, a saber, universitarios frente a no universitarios. Una variable que se controló fue la procedencia, es decir, que todos los entrevistados fueran de la Gran Área Metropolitana<sup>6</sup>, preferiblemente de San José.

#### 3.2. El instrumento

Con base en el cuestionario producido por Calvo Shadid (2008), se elaboró un nuevo instrumento, algo similar al anterior, compuesto por once ítems, con diferentes confluencias de vocales para los fenómenos que se analizarán; estos fueron: *huevo*, *leopardo*, *huesos*, *construir*, *huecos*, *lácteos*, *hierba*, *hielo*, *aéreo*, *aeropuerto*,

*hierro*. Se elicitaron las once palabras a partir de preguntas directas cuyas respuestas se podían deducir de las ilustraciones que se presentaban, con el fin de obtener la palabra deseada sin que los investigadores la mencionaran previamente. De este modo, se evitaba ofrecer un modelo de pronunciación. Se consideró solo el estilo de habla cuidadoso, puesto que se trató de una serie de preguntas cuya respuesta era totalmente guiada por las ilustraciones. No se trató de muestras de habla en condiciones de laboratorio en ninguno de los casos. Sin embargo, se consideró que las muestras debían tomarse controlando de algún modo el ruido exterior, ya que iban a ser sometidas a análisis experimental y no se podía arriesgar que los resultados fueran modificados o sesgados por ruidos externos. Se consideraron los siguientes contextos: diptongo en posición inicial de palabra: *huevo*, *huesos*, *huecos*, *hierba*, *hielo*, *hierro*; hiato en posición átona: *leopardo lácteos*, *aéreo*, *aeropuerto*, y diptongo en posición interna de palabra: *construir*.

Primero, se llevó a cabo exitosamente una prueba piloto en la Universidad de Costa Rica, en la cual se entrevistaron tres estudiantes para probar el instrumento y los análisis de Praat.

#### 3.3. La muestra

La muestra final utilizada para la investigación constó de 33 sujetos. Se trató de una muestra no aleatoria por conveniencia, pues se debió controlar el ruido exterior y someter a los entrevistados a una grabación en lugar cerrado y libre de ruido hasta donde fuera posible (oficinas, casas de habitación y una institución educativa). Si bien no fue posible eliminar la totalidad del ruido exterior, se trató de que los segmentos en los que se articulaban los diptongos o confluencias de vocales, específicamente, no se hubieran dado simultáneamente con ruidos exteriores como motores o bocinas de vehículos, personas hablando, y otros. De tal modo, se distribuyó la muestra en 33 hablantes: 16 hombres y 17 mujeres; 18 personas con instrucción universitaria y 15 sin instrucción universitaria; 20 del grupo de edad de 19 a 35 años, y 13 del segundo grupo, de 36 años en adelante.

### 3.4. Análisis del espectro

Como contraparte del análisis fonético hecho por oído, se produjeron espectros de frecuencias de la propia señal acústica con grabaciones de alta calidad. La interpretación de estos espectros viene detallada más abajo. Como es usual, la producción de espectros se basa en estudiar no la señal acústica en sí (o sea, las fluctuaciones de presión atmosférica) sino su transformada de Fourier (es decir, el espectro de frecuencias) instantánea.

La transformada de Fourier es la integral sobre el tiempo de la señal acústica pero pesada por una función de la forma  $\sin$  (frecuencia x tiempo). Los gráficos típicos usados en fonética acústica provienen de ordenar temporalmente a lo largo de un eje horizontal las distintas transformadas de Fourier instantáneas. Para llevar a cabo este proceso, se empleó el software libre Praat, desarrollado por los fonetistas Paul Boersma y David Weenink de la Universidad de Amsterdam. Este software calcula, representa y manipula las transformadas de Fourier y está diseñado para su aprovechamiento lingüístico (por ejemplo, marcando tono, formantes y energías de manera efectiva).

Para cada uno de los 26 informantes, se procedió a efectuar la segmentación de cada

una de las 11 palabras de interés, de tal manera que cada una aparecía aislada en un archivo. Se subsegmentó, a su vez, el rasgo de interés (por ejemplo, /we/ en /we.kos/). Se crearon así un total de 572 archivos .wav de alta calidad. A continuación, se obtuvo una imagen de cada uno de los 572 segmentos mencionados en dos formatos: uno que muestra el rango 0 a 7 kHz en frecuencias, y otro que muestra el rango 0 a 15 kHz. Esto suma un total de 1144 imágenes.

Finalmente, y con el fin de tener un machote de comparación, se procedió a hacer un catálogo con los prototipos (grabados por la autora) de las variantes más cerradas y más abiertas de los rasgos sonoros en cuestión (e.g., /we/ como oclusiva, fricativa, aproximante, diptongo, hiato). El catálogo establece una correspondencia entre audio e imagen (tanto digital como impresa) y su respectiva documentación para hacerla amigable.

## 4. Los resultados

### 4.1. El análisis articulatorio general

El análisis articulatorio general dio como resultado lo siguiente, de acuerdo con las confluencias de vocales analizadas:

Diptongos<sup>7</sup> en posición inicial:

## CUADRO 1.

Distribución de la muestra por diptongos en posición inicial de palabra

Diptongo inicial	Con consonantización	Sin consonantización	Ítems eliminados	Total
huevo(s)	16 51,61%	15 48,38%	2 <sup>8</sup>	31 100%
huesos	10 30%	23 70%	0	33 100%
huecos	7 21,21%	26 78,79%	0	33 100%
Total #ue-	33 34,02%	64 65,98%	0	97 100%
hierba	6 18,75%	26 81,25%	1 <sup>9</sup>	32 100%
hielo	0	33 100%	0	33 100%
hierro	0	33 <sup>10</sup> 100%	0	33 100%
Total #ie-	6 6,12%	92 93,88%		98
TOTAL	39 20%	156 80%	3	195 100%

A partir de los resultados anteriores, se puede deducir que la consonantización en los diptongos en posición inicial no es un rasgo muy característico en el español de la muestra del Área Metropolitana costarricense en el habla cuidadosa, con excepción de *huevo(s)*, que presenta una frecuencia de consonantización de poco más del 50%, la consonantización va descendiendo paulatinamente: a 30% en *huesos*, a 21,21% en *huecos*, a 18,75% en *hierba* y a ninguna consonantización en *hielo* y en *hierro*. Ahora bien, si elimináramos las lexías *huevo(s)* y *hierba*, por ser las que más normalmente se utilizan con consonantes en posición inicial, se podría afirmar que en esta muestra, la consonantización sería muchísimo menor.

En un total de 195 diptongos articulados en posición inicial de palabra, presentaron consonantización 39 de ellos, un 20%, lo cual se podría considerar una frecuencia baja de

consonantización delante de los diptongos iniciales. Con esta muestra se ha corroborado que en su mayoría (80%), los diptongos iniciales se pronuncian sin consonantización en su posición inicial. Así pues, fonológicamente, podrían existir los diptongos en posición inicial en español, tal y como lo determinó Calvo Shadid (2008:135): “La investigadora, a la luz de los datos, supone que la interpretación fonológica más adecuada sigue siendo la tradicional, propuesta por Navarro Tomás, la cual expone que existen diptongos en posición inicial absoluta y que estos se representan por los fonemas /i/ /u/ y la vocal – alta del núcleo.”

En general, se presenta un porcentaje muy bajo de consonantizaciones ante *hie-* (6,12%), frente a las que se articularon ante *hue-*, (34,02%). Por el contrario, la escasa consonantización ante *hie-* también confirma los hallazgos presentados por Calvo Shadid (2008:135): “En



la pronunciación de los diptongos iniciales *ie*, se ha observado, en un 100% de los casos, la realización fonética de [j] aproximante o semiconsonántica, sin ningún tipo de obstrucción o consonantización.” Es interesante señalar que incluso en dos oportunidades se articularon diptongos iniciales en sílabas diferentes, a saber, [i.ʔe.io], lo cual corrobora la pronunciación “esmerada” por influencia de la instrucción y de la escritura.

Por otra parte, y dado que hay cierto porcentaje de consonantización delante de

los diptongos iniciados en *ue*, y aunque no se comparte la interpretación fonológica de Alarcos Llorach respecto de su secuencia de fonemas /gu/ en posición inicial de palabra (difonemática), sí se comparte su propuesta de un refuerzo consonántico solo para [w], [gw] o [ɣw], en posición inicial de palabra, lo cual se corrobora en la muestra estudiada con un rango de reforzamiento entre el 20% y el 50% aproximadamente.

Sobre los hiatos<sup>11</sup> en posición inacentuada, los resultados fueron los siguientes:

## CUADRO 2

Distribución de la muestra por hiatos en posición inacentuada y por diptongo en posición interna de palabra

Hiatos inacentuados	Hiato	Diptongo	Ítemes eliminados	Total
leopardo	4 12,50%	28 87,50%	1 <sup>12</sup>	32 100%
lácteos	7 23,08%	26 <sup>13</sup> 76,92%	0	33 100%
aéreo	8 25,80%	23 <sup>14</sup> 74,20%	2 <sup>15</sup>	31 100%
aeropuerto	4 12,90%	27 87,10%	2 <sup>16</sup>	31 100%
Total	23 5,52%	104 81,88%	5	127 100%
Diptongo interno	Hiato	Diptongo	Ítemes eliminados	Total
construir	25 78,12%	7 21,88%	1	32 100%

A pesar de que las confluencias de vocales *eo* y *ae* se consideran hiatos en español desde “el punto de vista normativo” (ver nota 7), una gran mayoría de estas confluencias se articulan como diptongo con algunos ascensos vocálicos, cuatro en [ˈlak.tjos] ‘lácteos’ y cuatro en [a.ʔe.rjo]<sup>17</sup>; sin embargo, en su mayoría se trató de la articulación de la *e* como deslizada, tanto en *eo* como en *ae*. En general, los resultados no muestran una alta frecuencia de pronunciación de hiatos (23% en total), pero en todos los ítemes

estudiados se presenta aunque sea en porcentaje bajo. En *lácteos* es en la que más se observan hiatos, a saber, 23%.

Si bien, se comparte con Quesada Pacheco su afirmación: (2002: 66): “[...] parece ser un rasgo general del continente americano la eliminación de los hiatos, llamada también deshiatización [...]” no se comparte con el autor su afirmación sobre los grupos vocálicos siguientes: “/ea, ee, eo, oa, oe/ se convierten en diptongos crecientes”. Al menos en los

casos estudiados para /eo/ y /ae/ (este último no señalado por Quesada Pacheco) se registran solo diez ascensos o cierres vocálicos (cuatro para *lácteos* como [ˈlak.tjos] y seis para aéreo como [a.ˈe.rjo]) entre los 104 diptongos que se pronunciaron, es decir, solamente el 9,41% de los casos se realizaron como diptongos crecientes, por lo cual el rasgo en las secuencias vocálicas indicadas no pareciera ser un hecho.

Por último, dado que en Calvo Shadid (2008) los diptongos en posición interna de palabra no observaron variación, se utilizó en este caso un solo ítem, a saber, *construir*, que no se contempló en el estudio anterior. Los resultados son los siguientes:

Por otro lado, a diferencia de los casos anteriores, la palabra *construir* presenta en su articulación una mayoría de hiatos que de diptongos (casi en proporción 80% / 20%). Es probable que se haya dado así por el tipo de confluencia vocálica, a saber, *ui*. La situación de esta secuencia de vocales en español se expone en el *Diccionario Panhispánico de Dudas* (2005: 339), en el cual se señala:

En amplias zonas de América es muy frecuente que se pronuncien como diptongos secuencias formadas por una vocal cerrada átona y una vocal abierta tónica, o por dos vocales cerradas distintas; estas mismas secuencias, por el contrario, se pronuncian mayoritariamente como hiatos en el español de España y de algunos países americanos, por ejemplo, la Argentina o el Ecuador. Así, en España y parte de América, la palabra *guión* se pronuncia como bisílaba ([gi - ón]), debido al hiato, mientras que en otras áreas americanas se pronuncia como monosílaba ([gión]), debido al diptongo; lo mismo sucede con *jesuita*, que vacila, según las zonas, entre el hiato (je - su - i - ta) y el diptongo (je - suí - ta).

Así pues, contrario a lo que señala la cita, se podría afirmar que en la muestra estudiada, también hay una tendencia a pronunciar como hiato la secuencia *ui* (dos vocales cerradas distintas) como en *construir*, al menos en el habla cuidadosa.

#### 4.2. Análisis de datos por sexo

A pesar de las críticas que se han esgrimido en torno a la sociolingüística por

darle protagonismo a las diferencias de sexo o género y a las afirmaciones poco fundamentadas como las de que el habla de las mujeres es más “conservadora, insegura, flexible, solidaria y expresiva”, y la de los hombres más “independiente, competitiva y jerárquica”<sup>18</sup>, se ha tomado en cuenta en la investigación esta variable, con el fin de dar paso a una comprobación de hechos lingüísticos relativos a la pronunciación, sin emitir o aventurar criterios subjetivos o impresionistas. Los datos se muestran a continuación.

Los datos por sexo indican que los hombres, por una leve diferencia de solo el 3% que podría variar en una muestra mayor, presentan más articulaciones de una consonante delante del diptongo inicial, en especial, en los casos de *huevo* y *huesos*. Diferencias mínimas como esta en una muestra pequeña no dicen mucho acerca de la tendencia a la consonantización de los diptongos en posición inicial de palabra por el sexo; por el contrario, en todos los otros casos, se mantiene la ausencia de consonantización para ambos grupos. Incluso, se observó que tanto un hombre como una mujer, en la palabra *hierro*, articularon como hiato la secuencia de las vocales.

Cuando se trata de de los hiatos en posición inacentuada, ambos grupos presentan mayoría de diptongaciones. Igualmente, las mujeres tienden a articular más hiatos en estos contextos, aunque la diferencia también es mínima en relación con los hombres (solo 3%) en casi todos los casos, con excepción de *lácteos*, palabra en la cual los hombres articulan más hiatos que las mujeres. Los ascensos vocálicos, tanto como las metátesis de *aeropuerto*, se dan en ambos grupos.

Por otro lado, los diptongos internos en *-ui-* *construir*, presentan resultados por sexo semejantes a los anteriores: en ambos sexos se articulan como hiatos en su mayor cantidad, en especial entre las mujeres, cuya frecuencia de articulación es mayor que en los hombres en casi un 10%. La tendencia de esta confluencia de vocales es, pues, hacia el hiato, y en este análisis se observa que lo es tanto en hombres como en mujeres.

### 4.3. Análisis de datos por grupos de edad

Los dos grupos de edad en los que se dividió la muestra se deben a que con gran frecuencia en los estudios lingüísticos se hace un cambio de grupo generacional a partir de los 36 años. Por el tamaño de la muestra, no se estimó conveniente dividirla en más de dos grupos etarios, a saber, adultos jóvenes y adultos.

En posición inicial de palabra, el segundo grupo de edad, de 36 años y más, muestra una mayoría de consonantizaciones ante diptongos en posición inicial respecto de la primera, con una diferencia del 10% aproximadamente, especialmente en *hierba*, cuyas consonantizaciones en el grupo de edad mayor muestran una diferencia de más de un 30% respecto de la generación más joven. El segundo rubro en el que hubo más consonantizaciones en el segundo grupo de edad, es en *huecos*, con casi un 20% de diferencia respecto del grupo más joven. En *huevo(s)*, la diferencia se invierte: el grupo de edad más joven presenta mayoría de consonantizaciones, con una diferencia de casi un 10% respecto del segundo grupo. Así pues, aunque se muestra un patrón general de mayor consonantización en la generación mayor, llama la atención la mayor frecuencia de consonantizaciones en el grupo menor en *huevos*, (10% más). Se podría aventurar aquí alguna hipótesis como la de la lexicalización del término con la consonante incorporada para ciertos usos, especialmente en la generación más joven, pero se requieren más datos para aclarar esta leve tendencia.

Respecto de los hiatos en posición inacentuada, lo más común en un 80% en general, es la diptongación. Las palabras cuya pronunciación más registra hiatos es *aéreo*, y se da en la población más joven, con más del 30% de frecuencia de uso, y luego, *leopardo* y *lácteos*, con la misma cantidad de apariciones en cada grupo de edad, en el primero y el segundo grupo, respectivamente. Los hiatos son en general un poco más usados por el primer grupo de edad, con casi el 22% de uso, pero la diferencia entre ambos grupos no es alta, sino de alrededor del 8%. En *aeropuerto*, los hiatos

son menos frecuentes en ambos grupos de edad, aunque estos se dan más en el primer grupo que en el segundo. Así pues, la tendencia general en los hiatos en posición inacentuada es a pronunciarlos como diptongos en ambos grupos, con una ligera tendencia a usar más hiatos en el primer grupo de edad.

El diptongo interno en *-ui-* presenta una tendencia, por el contrario, a la hiatización, principalmente en el primer grupo de edad (85,71%), pero su frecuencia tampoco es despreciable en el segundo grupo (50%). Hay una diferencia de un 35% de mayor uso de hiatos en la generación más joven. El diptongo en estos casos es poco frecuente en el primer grupo de edad, 14,29%, en comparación con el segundo, a saber, 41,67% %, con una diferencia del 27,38%. En este caso, la generación más joven es la que mayor tendencia presenta a usar hiatos en este contexto.

### 4.4. Análisis de datos por escolaridad

En esta variable, por el tamaño de la muestra no se utilizaron escalas de nivel escolar. Se atendió únicamente a dos grupos: universitarios y no universitarios. Dentro de los no universitarios, se tomaron desde personas con instrucción de mitad de primer grado de primaria, hasta secundaria completa, y los universitarios desde estudiantes de tercer año de carrera, hasta profesionales con licenciatura, maestría y doctorado.

Por escolaridad, se observa que el grupo no universitario tiende un poco más, en términos generales, a introducir una consonante antes del diptongo inicial en casi un 9% más que el grupo universitario, pues, excepto en *huecos*, se observa una mayoría de consonantización inicial en ese grupo. La palabra que más se consonantiza es *huevo*, seguida por *huesos*, luego por *hierba*. La que menos consonantización presenta, en general, es *huecos*.

Acerca de los hiatos en posición inacentuada, el grupo no universitario presenta un porcentaje menor de articulación de hiatos que el universitario, con una diferencia entre ambos de 6% que quizás podría o aumentar

o reducirse en una muestra mayor, por lo que no pareciera haber una diferencia importante en la pronunciación de los grupos de distinta escolaridad en este contexto. La gran mayoría de pronunciaciones como diptongos en ambos grupos es obvia. Solamente en un caso, en la palabra *leopardo*, el grupo no universitario muestra una leve mayoría de pronunciación de hiatos y, por el contrario, el grupo universitario lo muestra en *aéreo*, con muchas más realizaciones de hiatos que el grupo no universitario.

Por último, en el diptongo *construir*, la pronunciación de hiatos es ligeramente mayor entre los universitarios (casi un 10% de diferencia) que entre los no universitarios. La diferencia no es mayor, y es importante señalar de nuevo que, en términos generales, la gran mayoría en ambos grupos presenta la hiatización de este diptongo.

## 5. Los resultados del análisis espectral

### 5.1. Planteamiento

Después de haber utilizado las grabaciones de alta calidad para realizar un análisis articulatorio, surge naturalmente la pregunta de si puede establecerse algún tipo de correspondencia entre dicho análisis y los espectros visuales generados a partir de un ordenamiento temporal de las transformadas de Fourier (es decir, el espectro). Dicha correspondencia sería muy valiosa, pues presentaría dos representaciones del mismo fenómeno.

No es de esperar que la correspondencia entre rasgos fonéticos percibidos auditivamente y los espectros visuales sea directa o fácil. Al fin y al cabo, el sistema auditivo (oído y cerebro) es mucho más sofisticado que el visual. Armados con el catálogo descrito al final de la sección 3 como machote, uno puede adoptar dos puntos de vista.

En primer lugar, uno podría usar el conocimiento visual para confirmar experimentalmente el análisis articulatorio. En este sentido veremos que hay una correlación positiva entre ambos análisis, y por tanto una corroboración, aunque modesta, del análisis articulatorio.

Por otra parte, y tomando en cuenta que el oído es, como se dijo, más sofisticado que la vista (incluso con la ayuda de software visual), uno podría más bien usar el análisis articulatorio como patrón de calibración y estudiar qué tan buena es la determinación visual de rasgos sonoros. Como veremos, es este punto de vista el que más se adecúa a los resultados.

### 5.2. Establecimiento de parámetros visuales

Se pasa a continuación a describir el procedimiento específico y sus resultados.

Se tomaron los datos correspondientes a diptongo en posición inicial (es decir, los casos “huevos”, “huesos”, “huecos”, “hierba”, “hielo” y “hierro”), y se establecieron tablas que vinculan características visuales (físicas) con rasgos sonoros.

CUADRO 3

Correspondencia entre rasgos acústicos (columnas) y características visuales (filas) para el caso de [we] inicial

	Oclusivo	Fricativo	Aproximante	Diptongo
Borde abrupto	Sí	No	No	No
Concavidad	Sí	Sí	Sí	No
Turbulencia previa	Sí	Sí	Leve	No
Formante previo	N/A	Sí	Muy leve	No

## CUADRO 4

Correspondencia entre rasgos acústicos (columnas) y características visuales (filas) para el caso de [we] inicial

	Africado	Fricativo	Diptongo
Borde abrupto	Sí	No	No
Turbulencia previa	No	Sí	No
Formante previo	Sí	Sí	No
Inclinación de formantes	Poca o nula	Intermedia	Mucha

A continuación, se procede a explicar los cuadros:

Se concentra la atención en cuatro características visuales en cada caso. Es muy importante que este número no sea demasiado alto, pues de lo contrario el algoritmo no sería viable. Estos cuadros están pensados para usarse con espectros del rango 0 a 15 kHz, que resultaron ser los más útiles.

“Borde abrupto”, en ambos cuadros, se refiere a un rasgo fino oscuro vertical en las altas frecuencias (definidas acá como frecuencias en la mitad superior, o sea 7,5 kHz a 15 kHz) y previo temporalmente a los formantes vocálicos.

“Turbulencia previa”, en ambos cuadros, se refiere a un rasgo menos fino que el anterior, más difuso, menos oscuro y menos prominente, también en las altas frecuencias y previo temporalmente a los formantes. (La palabra “turbulencia” apunta al origen físico de la apariencia del espectro).

“Formante previo”, en ambos cuadros, se refiere a la existencia de alguno de los formantes vocálicos, en las bajas frecuencias, con existencia *prematura*, es decir temporalmente previo al conjunto de formantes vocálicos en su totalidad.

“Concavidad”, en el cuadro 14, se refiere a la concavidad matemática hacia arriba del segundo formante vocálico. “Sí” significa que la concavidad de este formante es hacia arriba (como en el trazo “U”) o nula, en tanto que “No” indica una concavidad hacia abajo. La

concavidad hacia arriba suele estar acompañada de un boquete en su inicio.

“Inclinación de formantes”, en el cuadro 4, se refiere a la inclinación de los formantes vocálicos con respecto a la horizontal. Entre más se inclinen estos formantes, más abierta es la articulación.

Nótese que los cuadros no incluyen los hiatos. Esto es debido a dos motivos. En primer lugar, es muy difícil diferenciar diptongo de hiato visualmente. En segundo lugar, esta distinción no fue crucial en el análisis articulatorio. Cabe mencionar también que, en la categoría “fricativo” del cuadro 15, se agrupan dos articulaciones discutidas en el catálogo (una más cerrada, otra más abierta, pero ninguna deslizada).

### 5.3. Correlación entre rasgos sonoros y características visuales

A continuación, y armados con estos cuadros, se procedió a predecir el rasgo sonoro a partir del conocimiento visual del espectro, y a compararlo con el resultado del análisis articulatorio como referencia.

Se tomó como acierto si ambos datos caían en la misma categoría de consonantización o no-consonantización. Es decir, si el análisis visual predecía “diptongo” y el articulatorio decía “oclusivo”, se consideraba un desacierto, pero si el análisis visual predecía “fricativo” y el articulatorio decía “oclusivo”, se consideraba un acierto al ser ambos casos consonantizaciones.

El método visual, que no requiere entrenamiento auditivo, acertó en un 67% de las veces. (Si se distinguía entre los tipos de consonantización---por ejemplo, marcando como desacierto si el análisis visual precedía “fricativo” y el articulatorio decía “oclusivo”, entonces el porcentaje de aciertos cae ligeramente a un 64%).

Esta estadística no es tan baja. Sin embargo, si tomamos solamente los casos de consonantización, hacemos notar que la probabilidad de una identificación correcta, o sea de que el caso no se tome por diptongo o hiato, es apenas del 45%.

## 6. Conclusiones

A partir de los resultados anteriores, se llega a concluir lo siguiente:

Los datos por sexo indican, prácticamente en todos los casos, que hay pocas diferencias en la pronunciación de hombres y mujeres, con excepción quizás de *construir*, cuya pronunciación como hiato en las mujeres se eleva a un 10% más que en los hombres.

La consonantización en los diptongos en posición inicial no pareciera ser un rasgo característico del español actual costarricense en el habla cuidada, de esta muestra, con excepción de *huevo(s)*, que presenta una frecuencia de consonantización de más del 50%. Los diptongos iniciales se pronuncian sin consonantización en su posición inicial casi en un 80%. De ser por esta muestra, se podría afirmar que fonológicamente existen diptongos en posición inicial en español, tal y como lo propuso Calvo Shadid (2008:135). Se presenta un porcentaje mucho menor de consonantizaciones ante *hie-* que ante *hue-*, lo cual reafirma la posición de la autora en su texto del 2008:135: “aunque no se comparte la interpretación fonológica de Alarcos Llorach respecto de su secuencia de fonemas /gu/ en posición inicial de palabra, sí se comparte su propuesta de un refuerzo consonántico solo para [w], [gw] o [yw], en posición inicial de palabra.” Luego, sobre los hiatos en posición interna de la palabra, resultan no ser pronunciados como tales en la muestra estudiada; también en alrededor

de un 80% se articulan como diptongos<sup>19</sup>. Igualmente, el diptongo *construir*, se articula en casi un 80% como hiato y no como diptongo.

Los datos por edad indican que el segundo grupo, de 36 años y más, muestra mayoría de consonantizaciones en diptongos en posición inicial respecto de la primera en un 10%. Llama la atención que solo en *huevos*, la primera generación muestra mayor cantidad de consonantizaciones en un 10% que el grupo de edad mayor.

Los datos por escolaridad muestran que el grupo no universitario tiende a introducir una consonante antes del diptongo inicial en un 9% más que el grupo universitario. En el diptongo *construir*, la mayoría, en ambos grupos presenta la hiatización de este diptongo. La pronunciación de hiatos es ligeramente mayor entre los universitarios (casi un 10% de diferencia) que entre los no universitarios.

Por su parte, los datos experimentales arrojan una verificación modesta del análisis articulatorio. La correlación matemática no es alta, pero al menos es positiva (+0,15). La baja correlación matemática obtenida puede ser algo sorpresiva, sobre todo si se toma en cuenta lo concienzudo del procedimiento y la alta calidad de las grabaciones y de la producción de los espectros. Es sumamente interesante el hecho de que haya espectros muy distintos que se interpretan por el cerebro como el mismo rasgo sonoro. Esto sugiere, más que tratar de pulir el procedimiento de reconocimiento visual para obtener un mayor acuerdo entre métodos visual y articulatorio, que los resultados de esta parte de la investigación podrían usarse para iniciar un estudio cognitivo más profundo sobre las condiciones en las cuales una gran variedad de características físicas son percibidas por el sistema auditivo como la misma categoría de rasgo sonoro.

De tales resultados y conclusiones se desprende que los supuestos teóricos han sido todos comprobados en esta muestra de habla, a saber:

Los segmentos [w] y [j] no forman núcleo silábico con la vocal del diptongo, puesto que las

deslizadas en posición inicial de palabra pueden ocupar el ataque de la sílaba, tal y como se observa en su consonantización, específicamente en *huevos*, *huesos*, *huecos* y *hierba*.

Se supone que los diptongos en posición inicial de palabra presentan fonológicamente una vocal, a saber, /i/, /u/ la cual se desliza por contacto con la vocal plena y, en un proceso de resilabificación, la deslizada se puede convertir en una consonante prenuclear.

Las consonantizaciones en los diptongos en español no son tan frecuentes como lo proponía Navarro Tomás (1918). Sin embargo, se comparte con el autor la idea de que en la pronunciación cuidadosa y por influencia de la escritura, estas se articulan como semiconsonantes con alta frecuencia.

Los hiatos en posición inacentuada suelen realizarse como diptongos, esto es, en una misma sílaba, en el español de la Gran Área Metropolitana costarricense.

Algunos diptongos en posición interna de palabra suelen pronunciarse en dos moras silábicas, a saber, como dos núcleos silábicos independientes o hiatos.

## Recomendaciones

Se recomienda elaborar una investigación del tema de tipo cuantitativo en la Gran Área Metropolitana y otra a lo largo de todo el país para establecer la variación no solo sociolingüística, sino dialectológica de este y otros fenómenos del vocalismo en Costa Rica.

## Notas

1. Cf. El detallado marco teórico que Calvo Shadid (2008) presenta sobre el tema.
2. Tomado de Calvo Shadid (2008).
3. Alarcos simboliza con /y/ un fonema fricativo predorso prepalatal o palatal, probablemente /j/ en AFI.
4. Cfr. Navarro Tomás (1918).
5. En este caso, las deslizadas son semiconsonantes y semivocales que se dan en diptongos.
6. “La Gran Área Metropolitana (GAM) es la gran región metropolitana de Costa Rica que alberga a poco más de la mitad de la población [...] domina el sistema urbano de Costa Rica por su tamaño y por su localización en el centro de la nación [...] es una región multinuclear en donde las ciudades de San José, Alajuela, Cartago y Heredia son subcentros muy importantes [...] domina la mayoría de las actividades económicas del país [...] ocupa el 3,8% del territorio nacional, y concentra el 52,7% de la población total del país. [...] Históricamente la GAM ha concentrado la mayor cantidad de empleos del país.” Tomado de <http://ougam.ucr.ac.cr/index.php/la-gam>, la página del Observatorio Urbano de la Gran Área Metropolitana, de la Universidad de Costa Rica.
7. Se corrobora aquí que diptongo es la confluencia de vocales en una misma sílaba, y que esta confluencia no solo consta de una vocal alta y una no alta en la misma sílaba, sino que, como señala Cressey y se indica en el apartado 2.1, el español puede presentar cinco deslizadas que corresponden a cada una de las cinco vocales.
8. Se eliminaron por no aparecer en posición inicial: [uj.gwe.ɓo].
9. Se eliminó por no aparecer en posición inicial: [la.ʔjer.ba].
10. Se dieron tres pronunciaciones con hiato: [i.ˈe.io].
11. Se corrobora aquí que hiato es la confluencia de vocales, cualesquiera que sean, en sílabas diferentes. Según el *Diccionario Panhispánico de Dudas* (2005:339): Es la secuencia de dos vocales que se pronuncian en sílabas distintas: *grú-a*, *pa-ís*, *ca-er*, *dis-cu-tí- ais*. Desde el punto de vista fonético, son hiatos las combinaciones de vocal abierta (*a*, *e*, *o*) átona + vocal cerrada (*i*, *u*) tónica: *raíz*, *laúd*, *reír*, *transeúnte*, *oír*; de vocal cerrada tónica + vocal abierta átona: *María*, *rie*, *frío*, *cacatúa*, *acentúe*, *búho*; de dos vocales iguales: *azahar*, *poseer*, *chiita*, *alcohol*; y de dos vocales abiertas distintas: *caer*, *aorta*, *teatro*, *etéreo*, *coágulo*, *poeta* (aunque, en el habla, la secuencia de dos vocales abiertas — especialmente cuando ninguna de ellas es tónica— puede articularse como diptongo, esta combinación vocálica se considera siempre hiato desde el punto de vista normativo).

12. No pudo contestar la pregunta. Cressey, William. 1978. *Spanish phonology and morphology: A generative view*. Washington, D.C.: Georgetown University Press.
13. Se dieron cuatro ascensos vocálico: [ˈlak.tjos].
14. Se dieron seis ascensos vocálicos: [a.ˈe.rjo].
15. No pudieron contestar la pregunta adecuadamente. D'Introno, Francesco, Enrique del Teso Martín y Rosemary Weston. 1995. *Fonética y fonología actual del español*. Madrid: Ediciones Cátedra.
16. Se eliminaron por metátesis: [a.rjo.ˈpwer.to].
17. Al respecto, señala el *Diccionario Panhispánico de dudas* (2005: 339): “Tanto en España como en América existe una tendencia antihiática muy marcada en el habla popular, lo que provoca que determinadas secuencias vocálicas que son hiatos en el habla culta se pronuncien como diptongos entre hablantes poco instruidos, con el consiguiente cierre en el timbre de la vocal átona; [...]”. Mosterín, Jesús. 1981. *La ortografía fonémica del español*. Madrid: Alianza Editorial, S.A.
18. Como señala Peñas Ibáñez (2009:1).
19. No coincide en estos casos la afirmación de Quesada Pacheco (2002: 66) para el español de América, de que los grupos “/ea, ee, eo, oa, oe/ se convierten en diptongos crecientes”. En los casos estudiados para /eo/ y /ae/ (este último no señalado por Quesada Pacheco) se registran solo ocho ascensos o cierres vocálicos (cuatro para *lácteos* como [ˈlak.tjos] y cuatro para aéreo como [a.ˈe.rjo]) entre los 85 diptongos que se pronunciaron; solamente el 9,41%, por lo cual el rasgo en las secuencias vocálicas indicadas no pareciera ser característico del español de la GAM en el habla cuidadosa. Navarro Tomás, Tomás. 1918/61. *Manual de pronunciación española*. Madrid: Consejo Superior de Investigaciones Científicas.
- Navarro Tomás, Tomás. 1946/66. *Estudios de fonología española*. New York: Las Americas Publishing Company.
- Núñez Cedeño, Rafael y Alfonso Morales-Front. 1999. *Fonología generativa contemporánea de la lengua española*. Washington, D.C.: Georgetown University Press.
- Peñas Ibáñez, María Azucena. 2009. Semántica del discurso: La variable género. Recuperado de <http://es.scribd.com/doc/33138210/Semantica-del-discurso%C2%A8la-variable-genero>. Consulta: 30/01/2012.

## Bibliografía

- Alarcos Llorach, Emilio. 1961/83. *Fonología española*. Madrid: Editorial Gredos.
- Calvo Shadid, Annette. 2008. “Las semiconsonantes y semivocales en los diptongos del español: propuesta de análisis fonológico”. En: *Revista de Filología y Lingüística de la Universidad de Costa Rica XXXIV* (2): 107-142.
- Quesada Pacheco, Miguel Ángel. 2002. *El español de América*. Cartago: Editorial Tecnológica.
- Real Academia Española de la Lengua/Asociación de Academias de la Lengua Española. 2005. *Diccionario Panhispánico de Dudas*.

