Estudio fonético-acústico de los segmentos que realizan las secuencias grafémicas <hi> y <hu> seguidas de vocal en el castellano hablado en Chile*

Acoustic-Phonetic Study of Segments Using Orthographic Sequences <hi> and <hu> Followed by Vowel in Chilean Spanish

Lizbeth Gómez Aliste**, Mauricio Figueroa Candia***, Gastón Salamanca Gutiérrez****

RESUMEN

Este estudio entrega la primera evidencia destinada a determinar la naturaleza fonética de los sonidos articulados en las secuencias ortográficas <hi> y <hu> más vocal para el castellano hablado en Chile. Para ello, se anotaron instancias de estas secuencias y se midieron algunas de sus propiedades acústicas. Luego, se compararon los datos provenientes de las secuencias en estudio mediante modelos generalizados lineales mixtos con los de dip-tongos crecientes (por ejemplo [je] y [we]) y también con sílabas de tipo CV (por ejemplo [dʒe] y [ge]), con el fin de determinar si el primer elemento de las secuencias <hi> y <hu> más vocal tenía más parecido a unidades de tipo vocálico o de tipo consonántico. Los resultados sugieren que el primer elemento de las secuencias ortográficas tiene propiedades acústicas más similares a las de sílabas de tipo CV, lo que presenta importantes desafíos para

Palabras clave:
<hi> más vocal,
<hu> más vocal,
vocoide, contoide,
castellano chileno.

* Trabajo desarrollado en el marco del proyecto Fondecyt 1161466, titulado “Realizaciones fónicas segmentales en habla subestándar: una contribución a la construcción de un inventario inclusivo y actualizado de fonemas y alófonos del español de Chile”.

** Chilena. Magíster en Lingüística Aplicada de la Universidad de Concepción. Académica de la Universidad del BíoBío, Concepción, Chile. lizgomez@udec.cl

*** Chileno. PhD in Speech, Hearing and Phonetic Sciences del University College London. Académico de la Universidad de Concepción, Concepción, Chile. maufigueroa@udec.cl

**** Chileno. Doctor en Lingüística de la Universidad de Concepción. Académico de la Universidad de Concepción, Concepción, Chile. gsalaman@udec.cl

la interpretación de esta sección del sistema fonético y fonológico del castellano.

ABSTRACT

This study provides first evidence to describe the phonetic nature of the sounds articulated in the orthographic sequences <hi> and <hu> followed by vowel in Chilean Spanish. In order to do this, instances of these sequences were observed, and some of their acoustic properties were measured. The acoustic data obtained from the orthographic sequences was compared using generalized linear mixed models to those of Spanish opening diphthongs (e.g., [je] and [we]) and of CV syllables (e.g., [dʒe] and [ge]), in order to determine whether the first element of the <hi/hu> sequences is more similar to vocalic or consonantal units. The results suggest that the <hi> and <hu> orthographic sequences are more closely related, at least acoustically, to CV syllables, which poses interesting challenges for the interpretation of this section of the Spanish phonetic and phonemic systems.

Keywords: <hi> followed by vowel, <hu> followed by vowel, vocoid, contoid, Chilean Spanish.

1. Introducción

Los segmentos deslizados [j] y [w] (también llamados glides, o semi-consonantes) han formado parte de extensos debates que aún hasta hoy no se encuentran completamente zanjados (Bowen & Stockwell, 1955; Harris & Kaisse, 1999). En efecto, aunque diversos autores concuerdan en cuáles son los procesos fonéticos y fonológicos que ocurren en los diptongos del español, aún no se ha logrado aclarar de manera convincente cuáles son las realizaciones fonéticas que ocurren cuando estas glides van precedidas del grafema <h>, ni tampoco ha sido posible proponer una representación fonológica estable para estas secuencias. Este estudio se propuso contribuir a avanzar en el conocimiento de algunas de las propiedades acústicas de las secuencias conformadas por <hi> seguida de vocal y <hu> seguida de vocal, a través del uso de metodologías de la fonética acústica y de la estadística inferencial, con la finalidad de proveer insumos para que, en futuras incursiones investigativas, permitan resolver los problemas de orden fonológico que suscitan estas secuencias. Con la finalidad de poner en contexto la temática del presente estudio, entregaremos algunos antecedentes históricos para luego sintetizar los principales precedentes investigativos.

1.1. Antecedentes históricos respecto del uso del grafema “h” en español

Como es sabido, el grafema <h> no tiene representación fonética. En palabras de Navarro Tomás (1918), “la <h> no representa la aspiración laríngea que aparece en otros idiomas y que en otro tiempo parece haber existido también en nuestra lengua; la <h> ortográfica es actualmente en nuestra escritura una letra muda sin ningún valor fónico” (p. 59).

Ahora bien, aunque existe pleno consenso en que la <h> grafémica no tiene actualmente un sonido asociado, es razonable cuestionarse el porqué de su presencia, específicamente ante diptongo. Frente a esta duda, Hualde (2014) señala que su fundamento es puramente ortográfico e histórico, debido a la necesidad que existió en el español de distinguir entre palabras que comenzaban con un diptongo de las que resistieron los procesos evolutivos de diptongación al inicio de palabra, y debido a la ambigüedad que existía en la interpretación de los grafemas

<u> y <v>, antes de que fuesen letras diferentes; por ejemplo, algo escrito como “*uelo* [...] podría ser interpretado como ‘velo’ o como ‘huelo’. La <h> indica que la letra que sigue no puede ser una consonante” (Hualde, 2014, p. 305). Esta idea es compartida por la Real Academia Española (RAE), la que asegura la importancia del uso de la <h> en “palabras que empiecen por *ue* procedente de *o* latina: huevo, huérfano. Se justifica esta *h* para no equivocar la *v* con la *u* en la lectura y, así, en lugar de *uevo* no se leyese *vevo*, y *veso* en lugar de *hueso*” (RAE, 1754, p. 46).

Como consecuencia de estos criterios ortográficos, en español no existen diptongos crecientes en inicio de sílaba que no vayan precedidos por la <h> gráfica —excepto en muy escasas realizaciones de algunos préstamos como ocurre con las palabras “ion”, “iodo”, “iota” y sus derivados—, de ahí que sea más curioso aún que si dicha consonante nació por motivos de orden práctico, los diptongos que las contienen no sean exclusivamente vocálicos (o semivocálicos) como lo sugieren los precedentes de investigación que se revisarán a continuación.

1.2. Interpretación de los segmentos deslizados del español: precedentes

Desde estudios como el de Navarro Tomás (1918), hasta el de Hualde (2014), se ha propuesto que para los segmentos deslizados existen dos realizaciones distintas: semiconsonánticas [j] y [w] para los diptongos crecientes y semivocálicas [i] y [u] para los diptongos decrecientes. Aunque existen algunas discrepancias respecto de si es relevante representar estas unidades de manera diferente en los diptongos crecientes y decrecientes (Bowen & Stockwell, 1955), el hecho de que son alófonos no silábicos de /i/ y /u/ es generalmente una asunción que no se somete a discusión. Sin embargo, una excepción notable se puede encontrar en palabras que comienzan con la secuencia ortográfica <hi> más vocal, como las palabras “hielo” y “hierba”, y en <hu> más vocal, como en “huevo” y “hueso”. Autores como Alarcos Llorach (1961), Cepeda (1991), D’Introno, Saporta (1956), Teso y Weston (1995) y Quilis y Fernández (1999) concuerdan en que para <hi> más vocal se realizaría un fono africado en posición inicial, después de consonante nasal y luego de lateral, y que este podría alternar con una realización “fricativa” en los mismos contextos (hoy diríamos “aproximante”), pero en un habla poco cuidada. En cuanto a la secuencia gráfica <hu> seguida

de vocal, se realizaría siempre una oclusión en contextos fonéticos de inicio de palabra y después de nasal, descartando la opción de instancias semiconsonánticas, las que de ocurrir solo se manifestarían en un habla “vulgar” (Navarro Tomás, 1918).

Martínez-Celdrán, Fernández-Planas y Carrera-Sabaté (2003) fueron los primeros en proponer que podrían ocurrir fonos aproximantes en dichas secuencias gráficas. Exponen así que para la secuencia <hi> seguida de vocal existe una variante africada con aproximantización [t̪ʰ] después de consonante nasal, lateral o pausa, como en el caso de [elj.t̪ʰ.ate] “el yate”, y una variante aproximante [j] para el resto de los contextos como en [mi.j̥.ate] “mi yate” y desestiman la existencia de una variante fricativa. La presencia de alófonos aproximantes en las secuencias gráficas en estudio ha sido mencionada más recientemente por autores como Aguilar y Salamanca (2013), Borland Delorme (2004), Sadowsky y Salamanca (2011), entre otros.

Por su parte, Obediente (2007) —al igual que la mayoría de los autores presentados con anterioridad—, asume que los diptongos crecientes en posición inicial de sílaba sufren un proceso de consonantización en palabras como “hielo” y “hueso”, pero acepta que es posible que en ocasiones los hablantes influidos por la ortografía, articulen glides o incluso una vocal en contextos fonéticos de inicio de palabra y después de [s]. No obstante, para el autor, no es más adecuado considerar los segmentos que nos ocupan como alófonos de /d̪ʰ/ o de /g/, sino que propone un tercer alófono para ambas vocales altas, quedando en [i], [j] y [j̥] para /i/ y [u], [w] y [w̥] para /u/, siendo el último fono el que corresponde para los casos de <hi> y <hu> más vocal, como lo muestran ejemplos como “hielo” [ˈjelo] y “hueso” [ˈw̥eso], dado que dichas realizaciones no serían ni completamente consonánticas ni completamente vocálicas, hecho que sería entendido de esa manera por los hablantes.

Un estudio realizado por Calvo (2008) con hablantes de Costa Rica arrojó como resultado un bajo grado de consonantización para el segmento que realiza la secuencia gráfica <hu> seguida de vocal en posición inicial de palabra, y nula realización consonántica para <hi> más vocal en el mismo contexto; es más, para los mismos hablantes, consonantizar esta secuencia fonológica es considerado incorrecto. Para la autora queda claro que en los diptongos iniciados por <ue> existe un refuerzo consonántico de [w], por lo que son posibles las re-

alizaciones [gw] o [ɣw], pero en los casos de <ie> solo sería posible la realización fonética de una aproximante [j̞] o semiconsonante [j], pero nunca como una consonante plena.

Aguilar y Salamanca (2013), en un trabajo realizado con hablantes profesionales chilenos de la ciudad de Concepción y centrado únicamente en la secuencia grafémica <hu> más vocal, exponen que habría cuatro fonos discriminables, de los cuales el aproximante fue el que presentó mayor frecuencia de realización y la semiconsonante la menor cantidad de registros, contrariamente a lo que se reportó para el habla de Costa Rica.

Para terminar, Hualde (2014) concuerda con otros autores en que ambas secuencias grafémicas sufren una consonantización y que en posición inicial y después de nasal realizarían sonidos más implosivos. Es interesante resaltar, sin embargo, que en casos específicos como la palabra “hiena” Hualde señala que los hablantes tenderían a pronunciar la palabra como [ˈj̞e.na] en vez de [ˈje.na], dada la similitud de “hiena” con la palabra “llena”, mismo fenómeno que se daría con otros conjuntos como “de hueso”, que promovería una pronunciación más frecuente de [de.ˈj̞e.so] en lugar [de.ˈye.so], por su cercanía con conjuntos como “de guasa”, aunque estos contrastes ocurrirían solamente después de pausa y en un habla cuidada.

Como puede observarse, no existe consenso respecto de qué tipo de unidades fonéticas realizan los primeros elementos de las secuencias ortográficas <hi> y <hu> seguidas de vocal, puesto que los precedentes de investigación han detectado —mediante métodos diversos y para un conjunto de variedades dialectales del español, es cierto— un rango de realizaciones que va desde las más vocálicas hasta las más consonánticas; tampoco existe consenso en cuanto a cómo deben interpretarse fonológicamente estas realizaciones.

A continuación se expondrán los métodos que se han seguido en el presente estudio para obtener evidencia empírica que contribuya a caracterizar la naturaleza acústica de las realizaciones de <hi> y de <hu> seguidas de vocal, y los resultados de la aplicación de estos métodos. Como se verá, además de proveerse un reporte acerca de cuáles realizaciones son las más típicas, se realizarán comparaciones entre las instancias de las secuencias en estudio con realizaciones de diptongos

crecientes y de sílabas CV, con la finalidad de determinar si las instancias de <hi> y <hu> seguidas de vocal observadas en el castellano chileno son más parecidas a segmentos vocálicos o consonánticos.

2. Métodos

2.1. Participantes

Para esta investigación se utilizó un corpus de habla en lectura en voz alta elicitado de cuatro mujeres y cuatro hombres de ocho ciudades de zonas del país estratégicamente seleccionadas para abarcar parte importante del territorio chileno (Iquique, La Serena, Valparaíso, Santiago, Concepción, Temuco, Coyhaique y Punta Arenas). En total, se contó con un muestreo de 64 sujetos, todos nacidos y criados en la zona en cuestión, con edades que fluctuaron entre los 18 y 35 años. El estudio se centró en hablantes que no hubieran terminado su educación escolar formal y que no poseyeran ningún otro tipo de estudio superior o capacitación especial.

2.2. Instrumento, grabaciones y anotación

La lectura en voz alta se realizó sobre la base de 28 oraciones breves, de las cuales 24 contenían una o más palabras con las secuencias gráficas <hi> seguida vocal y/o <hu> más vocal, en diferentes contextos fonéticos tales como: posición inicial absoluta, posición intervocálica, después de /l/, luego de /s/, después de /n/, después de /r/ y en posición inicial después de pausa (Tabla 1). Las oraciones también contenían numerosas instancias de los fonemas /dʒ/ y /g/ seguidos de vocal, y de /i/ y /u/ en diptongos crecientes, cuyas realizaciones fueron utilizadas como punto de referencia para determinar si las actuaciones de las secuencias ortográficas en estudio tenían mayor parecido con unidades consonánticas o vocálicas. Los hablantes fueron grabados en salas insonorizadas, usando micrófonos Sennheiser EW-152-G3 montados sobre la cabeza. Estas señales fueron transmitidas por una grabadora digital Tascam DR-40, configuradas para una frecuencia de muestreo de 44.100 Hz y con una profundidad de 24 bits, en formato mono WAV. Es necesario explicitar que solo se utilizaron grabaciones provenientes de lecturas en voz alta debido al poco control que se tiene normalmente sobre las muestras de habla espontánea, lo que puede dificultar la obtención de cantidades suficientes de datos y una mues-

tra balanceada para un objeto de estudio relativamente infrecuente, como es nuestro caso (Aguilar y Salamanca, 2013; Cifuentes Becerra y Salamanca Gutiérrez, 2012).

Tabla 1.
Enunciados elicitados a través de lectura en voz alta y contextos en que ocurren las secuencias <hu> y <hi> seguidas de vocal.

Enunciado	Contextos
1. En el encuentro deportivo entre Hu achipato y Everton, jugó el Hu aso Isla, quien marcó un tremendo golazo de chilena, mientras escabullía la marca. Pero el árbitro lo anuló, haciendo sonar hue camente el pito.	V__V l__ r__
2. Los invitados a la fiesta se comieron hasta los hues os del chanco. Después tomaron piscola con hi elo, mote con hues illos del hu erto y helado de postre. Armaron hartó bochinche.	s__ n__ l__
3. Llegada la década de los ochenta, en Rengo, mi tío Gustavo y mi abuelo Pancho comenzaron a administrar una fábrica de muebles, con un alto nivel de seguridad.	(NA)
4. En la FACH, vendieron aviones de guerra en diez millones de dólares. Con el dinero recaudado, pagaron los sueldos pendientes de los guardias en huel ga.	n__
5. Ya fueron recuperados varios de los objetos robados a la Junaeb, debido a la rápida labor de los policías en la investigación sobre las hu ellas halladas en Llanquih ue .	s__ V__V
6. Risas provocó la ignorancia de Israel Marchant cuando admitió que escribió al revés el nombre de la laguna Conguillío en la web de la municipalidad.	(NA)
7. “Voy a darle ocho años y un día tras el hi erro de la prisión”, dijo el juez de Petro hu é. Cholo Muñoz no lo podía creer.	l__ V__V
8. Gabriel Llanos, triste por tener que beber hi edra venenosa como sentencia de muerte, admitió con llanto que el delito lo cometió alrededor de las tres. Le comenzó a correr hi elo por la espalda.	r__
9. Sánchez, Gari y Edson Puch, rápidos en contraatacar al rival, supieron ganar la copa con gran rendimiento. Sacaron ronchas en la barra contraria.	(NA)
10. Hue le bien la ciudad después de un día bien lluvioso, ya que con el agua se va el smog.	#__
11. Tras la asamblea, ciertas propuestas hicieron saltar chispas a los estudiantes de la Confech. Recibieron la oferta como soluciones parche.	- - -

12. Huevos revueltos con hierbas aromáticas ofrecían los chinos radicados en Huasco .	#__ n__
13. Beber whisky con hielo se puede hacer aprovechando los deshielos del sur chileno.	n__ s__
14. Chatear por la web, afirmaba Blanca, es mejor que patinar en el hielo .	l__
15. Triture el huachalomo con cuidado para no perder el jugo.	l__
16. Varios huasos hacen la pega y, al terminar, beben chelas para calmar la sed.	s__
17. Beto quedó atrapado en la red, pues ignoró la advertencia que había entre la hierba .	V__V
18. Ganaron los albos un partido absurdo. Huamachuco jugó mal, aunque estuvo cerca de dar vuelta el resultado.	\$__
19. Al llegar, hienas salvajes y otros animales nos atacan y nos hieren . Nosotros huimos en zigzag.	\$__ s__
20. El huemul y el pudú pigmeo están casi en extinción. Chile puede dejar huella positiva en las futuras generaciones si los protege.	l__ r__
21. “¡ Hielera me falta para la fiesta!”, gritó Magdalena tan fuerte que casi se le salen las amígdalas.	#__
22. “Si no me das de comer huevos , la hiel se me puede romper”, dijo Bob.	r__ V__V
23. “Yo quiero más hielo para el trago. Y pásame esas hierbas ”, dijo Lucho.	s__
24. “Morder huesos es lo que más le gusta a mi perro bulldog”, me dijo mi hermano, mientras el sabueso se zambullía en el lago.	r__
25. Organizar huelgas era la especialidad del dirigente.	r__
26. Deja que el agua hierva bien antes de poner hierbabuena en el mate.	V__V r__
27. Hiela mucho en la cordillera al llover fuerte.	#__
28. La profesora explicó a sus alumnos en qué consisten el hiato y el diptongo.	(NA)

Nota. El instrumento consideró 28 enunciados, cuya mayoría contiene las secuencias gráficas <hi> y <hu> más vocal y sus respectivos contextos fonéticos, además de numerosas instancias de diptongos crecientes y de los fonemas /dʒ/ y /g/. En la tabla se han destacado en negrita las instancias de <hi> y de <hu> seguida de vocal, y se incluye un listado de los contextos fonéticos en los que ocurren estas realizaciones.

Fuente: Elaboración propia.

El corpus, que ya había sido transliterado al nivel del enunciado, fue preprocesado usando el conjunto de paquetes de programación EasyAlign, para obtener automáticamente un segmentado y etiquetado del nivel de palabras, sílabas y límites de segmentos (Goldman, 2011). Todas las realizaciones de las secuencias y fonemas antedichos fueron luego identificadas y codificadas en TextGrids del programa PRAAT (Boersma y Weenink, 2018). La propuesta de etiquetado y segmentado para las unidades en estudio, realizada de manera preliminar por EasyAlign, fue revisada para todo el corpus a través de inspecciones auditivas y espectrográficas de las señales. Todas las instancias de diptongos crecientes ([ja], [je], [jo], [ju], [wa], [we], [wi], [wo]) presentes en palabras como “pierna” y “fuerte”, todas las instancias de las secuencias ortográficas <hi> y <hu> más vocal, y todas las instancias de /dʒ/ y /g/ más vocal fueron identificadas y codificadas en estratos de etiquetado separados, así como también información acerca del contexto fonético y sobre si las sílabas eran acentuadas o inacentuadas. En el caso de los fonemas /dʒ/ y /g/ seguidos de vocal, se etiquetaron y segmentaron solamente las realizaciones de los grafemas <y>, <ll> y <g>, que se dan en casos como “lluvia”, “yema” y “guardia”; se excluyeron, en cambio, las secuencias <ge> y <gi> pues se realizan fonéticamente como [xi̯e] y [xi̯i]. Nótese que en todos los casos se identificó el espacio de duración que ocupa el segmento de interés y la vocal siguiente, y se codificaron tanto los segmentos de interés como las vocales. En rigor, la unidad que medida fue la sílaba, aunque el interés estuviera focalizado siempre en el primer elemento.

Como se ha anticipado, la razón por la que se incluyeron instancias de diptongos crecientes y realizaciones de <y>, <ll> y <g> más vocal se debe a que constituyen extremos vocálicos y consonánticos, respectivamente, que servirán de punto de comparación para determinar si las realizaciones de <hi> más vocal y <hu> más vocal son más cercanas a instancias vocálicas o a consonánticas. Esta decisión asume que las ocurrencias de <hi> más vocal son contrastables con las de <y> y <ll> más vocal, y con las de diptongos crecientes con [j]. Dicho de otra manera, si se trata de determinar si las realizaciones de <hi> seguida de vocal son contoides o vocoides, es razonable asumir que instancias de estas secuencias con mayor grado de constricción se parecerán a la unidad fonológica que se representa ortográfica-

mente como <y> y <ll>, y que aquellas con menos constricción se parecerán a las de [j] de /i/; asumimos también que estos tres tipos de sonidos conforman un grupo de realizaciones con un primer elemento “anterior”. De la misma manera, se asumió aquí que las ocurrencias de <hu> más vocal son contrastables con instancias de /g/ y de diptongos crecientes con [w], y que estos tres grupos conforman un tipo de realizaciones con un primer elemento “posterior”.

Durante el análisis auditivo —tendiente a la anotación de las instancias de <hi> más vocal y <hu> más vocal—, el modo de articulación de las instancias de <hi> seguida de vocal se clasificó siempre como: fricativo, africado con distensión fricativa, africado con distensión aproximante, aproximante o como semiconsonante. En el caso de las secuencias de <hu> más vocal, los modos articulatorios utilizados para la clasificación fueron: oclusivo, fricativo, aproximante y semiconsonante.

La extracción de datos acústicos se realizó mediante el programa PRAAT, por medio de un script programado para esta investigación. Se obtuvieron mediciones de intensidad relativa para los tres tipos de sonidos mencionados (dipthongos crecientes, secuencias de <hi> o <hu> seguidas de vocal, e instancias de /dʒ/ y /g/ seguidas de vocal), medidos en los espacios de duración de cada realización y la vocal que la sigue, usando objetos de intensidad creados separadamente para cada hablante con valores por defecto. Las mediciones de diferencia de intensidad se calcularon restando el valor mínimo de intensidad detectado automáticamente en la sílaba a los máximos de intensidad detectados en la misma sílaba. Esta decisión asume que en general, para sílabas de este tipo, el mínimo de intensidad se encontrará cerca del comienzo de la sílaba, dada la mayor constricción articulatoria, y el máximo de intensidad cerca del punto central del núcleo vocálico, dado su menor grado de constricción articulatoria. Como quedará demostrado cuando se presenten los datos acústicos, esta asunción mostró ser verdadera para todas las realizaciones del corpus, pues no se registraron valores de diferencia de intensidad negativos. La decisión de utilizar diferencias de intensidad se justifica porque han sido interpretadas en la literatura como buenos correlatos de grado de tensión articulatoria (Carrasco, Hualde & Simonet, 2012; Hualde, Simonet, Shosted & Nadeu, 2010; Parrell, 2010); ade-

más, el procedimiento permite normalizar los valores de intensidad, a través de lo cual se están controlando posibles efectos, por ejemplo, de diferencias de intensidad entre participantes. Un ejemplo de las anotaciones y de los puntos de referencia de intensidad pueden apreciarse en la Figura 1.

Figura 1. Forma de onda, espectrograma, curva de intensidad y anotaciones en TextGrid de la expresión “varios huasos”.

Nota. Forma de onda, espectrograma, curva de intensidad y anotaciones en TextGrid de la expresión “varios huasos”, que contiene dos de las instancias de interés (un diptongo creciente y una secuencia de <hu> seguida de vocal). Los estratos del TextGrid, desde arriba hacia abajo, se codificaron de la siguiente manera: (1) transliteración de la expresión; (2) segmentos, transcritos en SAMPA (Wells, 1997); (3) clase, esto es, si la secuencia pertenece a un diptongo, a la secuencia ortográfica <hi/hu> seguida vocal o a una estructura CV con /dʒ/ o /g/ iniciales; (4) categoría, es decir, si el primer elemento de la secuencia es un segmento anterior o posterior; (5) acento; y (6) contexto fonético precedente. En la curva de intensidad de ambas secuencias, los valores mínimos y máximos de intensidad han sido identificados con círculos; el mínimo fue restado al máximo para obtener las diferencias de intensidad que se reportan en los resultados.

Fuente: Elaboración propia.

A riesgo de ser iterativos, conviene volver a mencionar que estas mediciones permitirán determinar si el grado de constricción de los fonos que realizan las secuencias grafémicas <hi> seguida de vocal y <hu> seguida de vocal son más cercanos a los de los diptongos crecientes o a las ocurrencias consonánticas de sílabas CV. La decisión respecto de si las realizaciones de las secuencias grafémicas en estudio constituyen un grupo independiente, o si corresponde interpretarlas más bien como instancias más cercanas a las de los diptongos crecientes o a las de realizaciones consonánticas se llevarán a cabo mediante herramientas de estadística inferencial, en ambiente R (R Development Core Team, 2008).

3. Análisis y resultados

Los procesos de segmentado y etiquetado arrojaron un total de 8.962 realizaciones. De estas, 3.790 son instancias de diptongos crecientes (42%), referidos de aquí en adelante como “diptongos”; 2.426 instancias fueron secuencias de <hi> o <hu> seguidas de vocal (27%), en adelante, “secuencias”; y 2.746 fueron realizaciones de /g/ o /dʒ/ seguidas por vocal (31%), a partir de ahora, “consonantes”. Estos tres niveles —diptongos, secuencias y consonantes— fueron agrupados bajo la variable categórica *clase*. Para el caso específico de las secuencias, la Tabla 2 resume los modos de articulación encontrados. Como se observa, las realizaciones aproximantes y los segmentos africados con aproximantización predominan en ambas categorías. Respecto de la zona de articulación, en total 4.648 instancias fueron categorizadas como pertenecientes al grupo “anterior” (52%), esto es, aquellas que comienzan con [j], <hi> o /dʒ/; y 4.314 instancias al grupo “posterior” (48%), esto es, aquellas iniciadas con [w], <hu> o /g/. Estos dos niveles —anterior y posterior— fueron agrupados bajo la variable categórica categoría. Por razones de espacio, las variables *acento* y *contexto fonético* no fueron incluidas en los análisis estadísticos que se reportarán a continuación, pero su importancia será abordada brevemente en la discusión.

Tabla 2.
Porcentajes de ocurrencia de los fonos que instancian el elemento inicial de las secuencias <hi> y <hu> más vocal.

Modo de articulación	Secuencias <hi>		Secuencias <hu>	
	%	AFI	%	AFI
Oclusivo	---	---	17,0%	[g]
Fricativo	0,8%	[j]	5,7%	[ɣ]
Africado 1 (distensión fricativa)	9,6%	[dʒ]	---	---
Africado 2 (distensión aproximante)	34,4%	[dʝ]	---	---
Aproximante	43,8%	[j]	76,7%	[ɣ]
Glides	11,4%	[j]	0,6%	[w]
Total	100%		100%	

Nota. Porcentajes de aparición y transcripción en Alfabeto Fonético Internacional (AFI) de los modos de articulación del primer elemento de las secuencias <hi> más vocal y <hu> más vocal, ordenados por grados aproximados de constricción, desde los que presentan mayor constricción (arriba) hasta los que presentan menor constricción articulatoria (abajo).

Fuente: Elaboración propia.

Los datos fueron importados a R, donde se construyó un modelo generalizado lineal mixto para evaluar los efectos de las variables clase y categoría (y su interacción) en la variable dependiente *diferencias de intensidad*. El modelo fue creado usando la función `lmer` del paquete de `lmerTest` (Kuznetsova, Brockhoff & Christensen, 2017). La variable *participantes* también fue incluida como un efecto al azar. Luego, se utilizó un procedimiento paso a paso para construir los modelos estadísticos: primero, se construyó un modelo nulo solo con la variable dependiente y el efecto aleatorio, para luego incluir las variables independientes y su interacción, una por una, reteniéndose las variables solamente cuando mejoraron significativamente el modelo, a juzgar por una función de análisis de varianza (en particular, la función ANOVA). Las tablas que contienen los efectos principales y las interacciones de cada modelo, con análisis de varianza de Tipo II, fueron producidas mediante la función ANOVA del paquete `car` (Fox & Weisberg, 2011), y usando la función `ranova` de `lmerTest`.

El modelo generalizado de efectos mixtos con mejor ajuste se presenta en la Tabla 3. De acuerdo con este modelo, hay un efecto principal de

clase ($\chi^2(2) = 3918,085, p < 0,001$) y categoría ($\chi^2(1) = 26,198, p < 0,001$), y una interacción significativa entre estas dos variables ($\chi^2(2) = 17,521, p < 0,001$). Los datos que explican estos efectos y la interacción pueden ser observados en el panel izquierdo de la Figura 2. Como el modelo sugiere, las instancias de diptongos ($\bar{x} = 6,36, \sigma = 4,4$) y consonantes ($\bar{x} = 15,83, \sigma = 10,02$) muestran una cantidad significativamente menor de diferencia de intensidad que las secuencias ortográficas en foco ($\bar{x} = 17,94, \sigma = 9,88$). Sin embargo, tanto el tamaño de los coeficientes como el de t —ver Tabla 3—sugieren que las diferencias entre diptongos y secuencias son considerablemente mayores que entre las consonantes y las secuencias. En otras palabras, las secuencias de <hi> más vocal y <hu> más vocal parecen estar más estrechamente relacionadas acústicamente con las consonantes que con los diptongos típicos. Respecto de la categoría, aunque se detectó un efecto principal—lo que sugiere diferencias entre las realizaciones anteriores y posteriores (consúltese el panel de la derecha de la Figura 2)—, este efecto se diluye cuando se tienen en cuenta los datos de clase (consúltese el panel de la izquierda de la Figura 2). Finalmente, respecto de la interacción entre clase y categoría, la diferencia entre los niveles anterior ($\bar{x} = 6,58, \sigma = 4,87$) y posterior ($\bar{x} = 5,95, \sigma = 3,3$) no es significativa cuando se comparan los diptongos con las secuencias, pero lo es cuando las instancias de las consonantes anteriores ($\bar{x} = 17, \sigma = 10,68$) y posteriores ($\bar{x} = 15,09, \sigma = 9,51$) se comparan con las secuencias ortográficas. Esto se debe, probablemente, a una mayor diferencia entre las realizaciones anteriores y posteriores en consonantes que en los diptongos.

Tabla 3.
Modelo generalizado de efectos mixtos para la variable dependiente diferencia de intensidad, incluyendo clase y categoría como efectos principales, y participantes como efecto aleatorio.

Efectos mixtos	Coeficientes	Error estándar	Valor t	Valor p
Intercepto	18,0009	0,3266	55,119	< 0,001
Clase				
Secuencia	(nivel de ref.)			
Diptongos	-11,4253	0,2829	-40,380	< 0,001
Consonantes	-1,0001	0,3362	-2,975	< 0,01
Categoría				
Anterior	(nivel de ref.)			

Posterior	-0,1196	0,3203	-0,373	= 0,70890
Clase *				
Categoría				
Secuencia/ Anterior	(nivel de ref.)			
Diptongos/ Posterior	-0,5078	0,4176	-1,216	= 0,22402
Consonantes/ Posterior	-1,7820	0,4444	-4,010	< 0,001
Efectos aleatorios	Verosimilitud logarítmica	Grados de libertad	Test de razón de verosimilitud	Valor p
Participantes	-31.428	1	320,68	< 0,001

Fuente: Elaboración propia.

Figura 2. Gráficos de cajas para los valores de diferencias de intensidad de los segmentos analizados.

Nota. Panel izquierdo: diagramas de diferencias de intensidad para la variable categoría (“dip” = diptongos; “sec” = secuencias; “cons” = consonantes) y clase (“a” = anterior; “p” = posterior). Panel derecho: diagrama de diferencia de intensidad para los datos agregados de la variable categórica categoría, pero separados por clase.

Fuente: Elaboración propia.

4. Discusión

Como se mencionó anteriormente, la evidencia acústica (Figura 2) y cualitativa (Tabla 2) muestra que las realizaciones de las secuencias <hi> y <hu> más vocal están más estrechamente relacionadas con consonantes como las que se encuentran en sílabas que comienzan con /dʒ/ y /g/. Respecto de <hu>, es relevante que a pesar de que para esta secuencia no aplican dos de los modos articulatorios que sí son instanciados en <hi>, queda claro que la consonantización es la norma y la articulación de glides la excepción. En efecto, si se suman las secuencias consonánticas, el conjunto de variantes no vocálicas arroja un 99,4% del total de las instancias de <hu>, y de ese porcentaje el 76,7% corresponde a realizaciones aproximantes. Esta tendencia no es tan evidente en el caso de <hi>, pues el porcentaje de realizaciones consonánticas es de un 88,6% (78,2% con algún componente aproximante); sin embargo, es claro también el predominio de las realizaciones con mayor constricción en lugar de las vocálicas. En este contexto, se comprueban diferencias importantes respecto de los resultados observados para el habla de Costa Rica (Calvo, 2008), y avenencias con el estudio presentado por Aguilar y Salamanca (2013).

En el caso de la evidencia acústica, nos parece que esta también muestra tendencias claras: al analizar el grado de constricción articulatoria de las instancias de <hi> y <hu> seguidas de vocal mediante un medio indirecto (las diferencias de intensidad) es posible observar y determinar, estadísticamente, que las realizaciones en foco presentan características más similares a sílabas de tipo CV que a diptongos crecientes. Es importante también observar que, además de que las secuencias <hi> y <hu> seguidas de vocal presentan importantes diferencias estadísticas respecto de las instancias de diptongos crecientes, las distribuciones observables en la Figura 2 no se traslapan, es decir, las sílabas con diptongo creciente son cuantitativa y cualitativamente diferentes de las instancias de <hi> y <hu> seguidas de vocal. Tomadas en su conjunto, la evidencia cualitativa y acústica sugieren que las instancias de las secuencias <hi> y <hu> seguida de vocal se comportan, en su mayoría, como sílabas de tipo CV, lo que tiene importantes consecuencias para nuestra comprensión del sistema fonológico castellano. Nótese también que por lo menos para efectos de su comportamiento articulatorio y acústico no se demostraron diferencias entre

las instancias de <hi> seguida de vocal y de <hu> seguida de vocal, lo que da cuenta de cierta homogeneidad o comportamiento de conjunto de las realizaciones de ambas secuencias ortográficas, a pesar de las diferencias que se observan en los modos de articulación que las representan y de los porcentajes de aparición que recibe cada uno de estos modos (Tabla 2).

Aunque este estudio no se propone arribar a una nueva propuesta de organización fonético-fonológica para las secuencias en estudio, parece ser el caso que las unidades representadas por <hi> y <hu> seguidas de vocal tienen tanto una naturaleza contoides como un comportamiento consonántico, dado que solo aparecen en el premargen silábico y que presentan un grado de constricción articulatoria similar al de otras consonantes comparables. Si su naturaleza y comportamiento es más bien consonántico, pero distinto al de los diptongos crecientes, entonces parece necesario revisitar la organización de esta sección del sistema fonético y fonológico del castellano chileno, y adoptar alguna de las siguientes tres hipótesis:

- que las instancias de <hi> y <hu> seguidas de vocal son alófonos contoides de /i/ y de /u/, los que podrían representarse tal vez como [j] y [w], para distinguirlos de las glides [j] y [w];
- que las instancias de <hi> y <hu> seguidas de vocal son alófonos vocálicos de /d̪/ y de /g/;
- que las instancias son alófonos de dos unidades fonológicas independientes, las que podrían representarse como /j/ y /w/.

En relación con algunas proyecciones, además de la reinterpretación del sistema fonético y fonológico que pueden contribuir a informar los datos acá presentados, aún resta por evaluar el efecto que pueden tener en las realizaciones de <hi> y <hu> seguidas de vocal otras variables de importancia tales como el acento léxico y el contexto fonético, ya que es probable que también posean un papel relevante. Para evidencia preliminar, véase la Figura 3, la que sugiere que la variable categórica acento no parece tener un efecto claro en las realizaciones de <hi> y de <hu> seguidas de vocal, y que la variable contexto fonético sí parece tenerlo, tal que las instancias articuladas luego de pausa silente presentan una constricción articulatoria superior al del resto de los contextos, incluso bastante mayor que las reali-

zaciones que se encuentran luego de consonante nasal. Otro ámbito que vale la pena explorar es si los resultados que se han observado aquí son parecidos o no a los que podrían obtenerse mediante otra tarea de elicitación, en la que los hablantes suelen utilizar un habla más cercana al habla espontánea.

Figura 3. Gráficos de cajas correspondientes a las diferencias de intensidad de las secuencias <hi> y <hu> organizadas por las variables categóricas acento y contexto fonético, respectivamente.

Nota. Panel izquierdo: gráfico de cajas de las diferencias de intensidad de las secuencias <hi> y <hu> seguidas de vocal organizadas por la variable categórica acento ("0" = no acentuada; "1" = acentuada). Panel derecho: gráfico de cajas de las diferencias de intensidad de las secuencias <hi> y <hu> más vocal organizadas por la variable categórica contexto fonético ("p" = después de pausa; "n" = después de nasal; "v" = después de vocal; "o" = otros contextos).

Fuente: Elaboración propia.

5. Conclusiones

El presente estudio acústico y estadístico acerca de la naturaleza de las realizaciones de las secuencias <hi> seguida de vocal y <hu> más vocal reveló en primer término que las mediciones de intensidad relativa presentan diferencias estadísticamente significativas respecto de las de los diptongos crecientes y de las de las consonantes en sílabas CV, pero que estas diferencias son considerablemente más importantes entre las secuencias y los diptongos, que entre las secuencias y las sílabas CV con un primer elemento consonántico. Dicho de manera más sencilla, es posible concluir que el grado de constricción de las realizaciones de las secuencias se acerca más al de secuencias de estructura CV.

Por otra parte, aunque fue posible observar pequeñas diferencias entre las instancias anteriores y posteriores, la variable clase no parece tener un rol particularmente importante cuando se trata de los resultados de <hi> y <hu> seguidas de vocal, lo que habla de un probable comportamiento de conjunto, a pesar de las importantes diferencias de punto articulatorio.

Toda esta evidencia, tomada en su conjunto, problematiza la interpretación actual que tenemos de cierta porción del sistema fonético y fonológico del castellano. Desde una perspectiva metodológica, también es posible concluir que las mediciones acústicas realizadas parecen codificar información relevante en cuanto a la variabilidad de los segmentos en estudio, y en las direcciones esperadas teóricamente.

Referencias bibliográficas

- Aguilar, E. y Salamanca, G. (2013). Fonos que realizan la secuencia grafémica “hu+vocal” en 17 sujetos de Concepción e hipótesis inicial sobre su organización fonológica. *Literatura y Lingüística*, 28, 193-214. <https://doi.org/10.4067/s0716-58112013000200011>
- Alarcos Llorach, E. (1961). *Fonología española*. Madrid: Gredos.
- Boersma, P. & Weenink, D. (2018). PRAAT: Doing phonetics by computer (programa computacional). Amsterdam: University of Amsterdam. Recuperado de <http://www.praat.org/>

- Borland Delorme, K. E. (2004). La variación y distribución alofónica en el habla culta de Santiago de Chile. *Onomázein*, 10, 103-115. Recuperado de <http://www.acuedi.org/ddata/4171.pdf>
- Bowen, J. & Stockwell, R. (1955). The phonemic interpretation of semi-vowels in Spanish. *Language*, 31(2), 236-240. <https://doi.org/10.2307/411039>
- Calvo, A. (2008). Las semiconsonantes y semivocales en los diptongos del español: propuesta de análisis fonológico. *Filología y Lingüística de la Universidad de Costa Rica*, 34(2), 107-142. <https://doi.org/10.15517/rfl.v34i2.1282>
- Carrasco, P., Hualde, J., & Simonet, M. (2012). Dialectal differences in Spanish voiced obstruent allophony: Costa Rican versus Iberian Spanish. *Phonetica*, 69(3), 149-179. <https://doi.org/10.1159/000345199>
- Cepeda, G. (1991). *Las consonantes de Valdivia*. Valdivia: Universidad Austral de Chile.
- Cifuentes Becerra, E. y Salamanca Gutiérrez, G. F. (2012). Desarrollo y aplicación piloto de un instrumento para recoger datos del nivel fonético-fonológico. *Literatura y Lingüística*, 25, 221-248. <https://doi.org/10.4067/s0716-58112012000100011>
- D'Introno, F., Teso, E., y Weston, R. (1995). *Fonética y fonología actual del español*. Madrid: Cátedra.
- Fox, J. & Weisberg, S. (2011). *An {R} companion to applied regression*. Thousand Oaks: Sage Publications.
- Goldman, J. (2011). *EasyAlign: An automatic phonetic alignment tool under PRAAT*. *Interspeech'11, 12th Annual Conference Of The International Speech Communication Association*. Recuperado de <https://archive-ouverte.unige.ch/unige:18188>
- Harris, J. W. & Kaisse, E. M. (1999). Palatal vowels, glides and obstruents in Argentinian Spanish. *Phonology*, 16(2), 117-190. <https://doi.org/10.1017/s0952675799003735>
- Hualde, J. (2014). *Los sonidos del español*. Nueva York: Cambridge University Press.
- Hualde, J., Simonet, M., Shosted, R., & Nadeu, M. (2010). *Quantifying Iberian spirantization: Acoustics and articulation*. Recuperado de https://pdfs.semanticscholar.org/1d72/a085285ac93d73a4561fbd76b7f2900a5f3e.pdf?_ga=2.31289405.365481033.1586095308-854027056.1566854036

- Kuznetsova, A., Brockhoff, P., & Christensen, R. (2017). lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1-26. <https://doi.org/10.18637/jss.v082.i13>
- Martínez-Celdrán, E., Fernández-Planas, A., & Carrera-Sabaté, J. (2003). Castilian Spanish. *Journal of the International Phonetic Association*, 33(2), 255-259. <https://doi.org/10.1017/S0025100303001373>
- Navarro Tomás, T. (1918). *Manual de pronunciación española*. Madrid: Consejo Superior de Investigaciones Científicas.
- Obediente, E. (2007). *Fonética y fonología*. Venezuela: Universidad de Los Andes.
- Parrell, B. (2010). Articulation from acoustics: Estimating constriction degree from the acoustic signal. *Journal of the Acoustical Society of America*, 128(4), 2289-2289. <https://doi.org/10.1121/1.3508033>
- Quilis, A. y Fernández, J. (1999). *Curso de fonética y fonología españolas: para estudiantes angloamericanos*. Madrid: Consejo Superior de Investigaciones Científicas.
- R Development Core Team. (2008). R: A language and environment for statistical computing (software computacional). Recuperado de <http://www.R-project.org>
- Real Academia Española, RAE. (1754). *Ortografía de la lengua española castellana compuesta por la Real Academia Española. Nueva edición corregida y aumentada*. Madrid: Autor.
- Sadowsky, S. y Salamanca, G. (2011). El inventario fonético del español de Chile: principios orientadores, inventario provisorio de consonantes y sistema de representación (AFI-CL). *Onomázein*, 24, 61-84. Recuperado de http://onomazein.letras.uc.cl/Articulos/24/3_Sadowsky.pdf
- Saporta, S. (1956). A note on Spanish semivowels. *Language*, 32(2), 287-290. <https://doi.org/10.2307/411006>
- Wells, J. C. (1997). SAMPA computer readable phonetic alphabet. En D. Gibbon, R. Moore, & R. Winski (Eds.), *Handbook of standards and resources for spoken language systems* (pp. 684-732). Berlin, New York: Mouton de Gruyter.