

#### 4.4.3.6.4. PROCESOS DE REDUCCIÓN FONÉTICA

A partir de los procesos de reducción fonética de los grupos vocálicos en habla conversacional es posible obtener datos adicionales para las combinaciones de dos vocales cerradas. Los hiatos con un porcentaje mayor de reducciones son los formados por la combinación de las vocales [i] e [u]: cuando [i] aparece en posición inicial, se reducen a diptongo el 8% de las combinaciones analizadas, mientras que si dicha posición está ocupada por [u], el 28% de los casos se manifiesta como un diptongo y el 1% como una vocal.

También en los diptongos, la combinación vocálica que presenta un mayor porcentaje de reducción es la de [i]+[u], reducida en un 92% de sus apariciones cuando [i] aparece en posición inicial del grupo; por el contrario, si [u] es el elemento inicial, el porcentaje de reducción desciende al 17%.

El grupo [ui], por tanto, es más resistente a la reducción que la combinación [iu]; sin embargo, el resultado de la reducción en [iu] es un elemento anterior, [i] o [e], el 94% de los casos, lo cual no permite interpretar que [u] sea fonéticamente más persistente que [i]. No podemos por otra parte utilizar los datos de reducción fonética como prueba de silabicidad, dado que si observamos el comportamiento de las reducciones de diptongos en las muestras procedentes del habla conversacional, no es el elemento que funciona fonológicamente como núcleo el que se mantiene un mayor número de veces, sino que se observa una tendencia a elidir el segundo elemento del grupo, independientemente de su naturaleza funcional.

### Resumen

El comportamiento duracional de los diptongos [iu], [ui] así como los valores de frecuencia de los formantes en los segmentos aportan algunos argumentos en favor de considerar a [u] como vocal silábica, independientemente de la posición que ocupe en el grupo. Por un lado, [ui] se comporta en cuanto a duración como un diptongo decreciente; por otro lado, [u] en el diptongo [iu] presenta valores similares a los de [u] vocal, resistiéndose a la influencia del

elemento palatal precedente. Asimismo, se ha podido observar una mayor duración de [u] frente a [i] cuando se promedian los valores del segmento en posición inicial y final del grupo. Tales observaciones coinciden con las descripciones clásicas de los diptongos del español en que no se reconoce la existencia del diptongo [iɥ]; sin embargo, la consideración de [u] como núcleo en [ui] no goza de la misma aceptación.

Sin embargo, determinar la silabicidad a partir de parámetros acústicos tales como la duración y la frecuencia de los formantes no es posible, dado que los correlatos acústicos de la misma no están previamente establecidos. Por otro lado, la función fonológica de un segmento y su comportamiento ante procesos de reducción fonética pueden entrar en conflicto: así, si bien basándonos en la categoría funcional de un segmento, se podría pensar que las vocales silábicas son las más resistentes a los procesos de reducción, en el plano fonético, hay otros condicionamientos -posición en el grupo- que determinan la manifestación de dichos procesos.

#### **4.4.3.7. PROCESOS DE ASIMILACIÓN**

En ambos conjuntos de datos, se aprecia un grado mayor de influencia de la consonante sobre las vocales contiguas, manifestada en forma de palatalización o velarización, que de la semiconsonante, la vocal en hiato y la vocal.

En cuanto a la comparación de la influencia ejercida por la semiconsonante y por la semivocal, aparece tanto en el corpus de palabras en frases marco como en el de la tarea del mapa un mayor influjo de la semiconsonante sobre la vocal que de la semivocal sobre la vocal. Dicha diferencia se da tanto en los procesos de palatalización como en los de velarización.

## Resumen

La comparación de las secuencias del corpus en dos situaciones comunicativas permite observar las modificaciones fonéticas que surgen con el cambio de estilo de habla. Hallamos diferencias en las realizaciones fonéticas -aparición de variantes distintas en contextos idénticos- así como en su distribución -divergencias en los contextos de aparición-. Por otro lado, la manifestación acústica de los grupos considerados -hiatos, diptongos, sílabas [j]V- [w]V- así como de los segmentos semivocálicos y semiconsonánticos presenta variaciones en ambos conjuntos de datos, tanto en el dominio temporal como en el de la frecuencia.



## 4.5. DISCUSIÓN

La búsqueda de invariantes en la señal acústica es un tema recurrente en los estudios fonéticos; a esta búsqueda, sin embargo, se le opone como principal problema la segmentación de las unidades lingüísticas.

*“The speech signal cannot be unambiguously segmented into temporally nonoverlapping chunks corresponding to linear sequences of phonemes, syllables and words (Segmentation issue)”*

Lindblom (1986)<sup>1</sup>

El problema de la segmentación se refiere a la división de la onda acústica continua en segmentos que se correspondan con las unidades fonéticas. Las dificultades aparecen cuando se observa que, en ocasiones, el número de segmentos acústicos excede al número de segmentos fonéticos, del mismo modo que el tipo de segmento acústico difiere según el contexto; por otro lado, no hay que olvidar que dentro de una unidad acústica, hay información sobre más de un elemento fonético debido al fenómeno de la coarticulación. La segmentación de la señal de habla por medio de líneas perpendiculares al eje del tiempo situadas allí donde se advierten cambios de la forma de onda da como resultado la constatación de la existencia de elementos influidos por el entorno. En este punto, se requiere una teoría que pueda explicar el proceso mediante el cual versiones articulatoria y acústicamente distintas se interpretan por parte del oyente como una única unidad fonética. Lindblom resume así el problema de la invariación fonética.

*“For a given language there seems to be no unique set of acoustic properties that will always be present in the production of a given unit (feature, phoneme, syllable) and that will reliably be found in all conceivable contexts. (Invariance issue)”*

<sup>2</sup>Lindblom (1986)

---

<sup>1</sup> pág. 495.

<sup>2</sup> pág. 495.

La cuestión fundamental en el tema de la invariación en los procesos fonéticos radica en proponer descripciones físicas de las unidades lingüísticas abstractas que posean la característica de permanecer invariables en el amplio número de contextos posibles de los actos comunicativos del lenguaje. Las teorías fonéticas cuyas formulaciones se basan en oposiciones binarias (Jakobson *et al.*, 1952; Chomsky-Halle, 1968) asumen la existencia de unidades invariables, discretas -rasgos fonéticos- que conforman el código del habla. Sin embargo, los enfoques sociolingüísticos cuestionan la existencia de tales universales, trasladando el énfasis a los factores de variación (Labov, 1972). En el terreno de la fisiología, algunos investigadores argumentan que la programación motriz del habla es dependiente del contexto aunque basada en valores invariables (Perkell, 1980), mientras que otros aportan una unidad en la que están incluidas las influencias contextuales y la producción se concibe como un proceso directo sin necesidad de ninguna transformación (Fowler *et al.*, 1980). Algunos modelos, finalmente, proclaman la necesidad de acabar con la oposición variación/invariación al no distinguir entre elementos fonológicos, invariables, e implementaciones fonéticas, variables: la variación y la invariación están en estrecha ligazón del mismo modo que la fonología se incorpora en los laboratorios de fonética (Kohler, 1986; Elman-Mc Clelland, 1986).

### *Teoría de la Invariación Acústica*

La teoría de la invariación acústica se basa en dos premisas. Por un lado, la invariación es directamente derivable de la señal acústica, en forma de propiedades integradas que no se relacionan con un componente individual sino que pueden incluir a varios. Estas propiedades están estrechamente relacionadas con los rasgos distintivos. Por otro lado, se asume que el sistema perceptivo está dotado de mecanismos detectores de tales propiedades (Stevens, 1975). El proceso de transformación de sonido en rasgos distintivos es directo, y aunque se reconoce la existencia de indicios dependientes del contexto sólo se les atribuye una función secundaria (Blumstein, 1980). Los indicios primarios son un conjunto de propiedades invariables de naturaleza innata, mientras que los parámetros secundarios se adquieren a partir de la interacción con el entorno lingüístico.

La naturaleza de las propiedades invariables y su relación con los segmentos fonéticos, según esta teoría, reflejan las restricciones del conjunto posible de sonidos por parte del mecanismo articulatorio (naturaleza cuántica de los sonidos) y del mecanismo perceptivo

(percepción categorial)<sup>3</sup>; también muestran la manera en que se definen las clases de sonidos, basados en rasgos distintivos. Existen propiedades acústicas que distinguen un segmento fonético de otro, además de unas propiedades generales que caracterizan a las unidades del habla (Blumstein-Stevens, 1979; Stevens, 1980). Blumstein-Stevens (1985) enfatizan el papel de los rasgos fonéticos en una teoría de la invariación acústica: las propiedades invariables corresponden a rasgos fonéticos, no a fonemas. Si un rasgo fonético cumple una función de contraste fonológico en una lengua, debe estar representado, por definición, por una propiedad acústica, que permanece inalterada a través de diferentes fuentes de variación, entre ellas el contexto vocálico, la posición silábica y el hablante. El papel de los rasgos fonéticos es definir la forma del sistema fonético-fonológico de la lengua y ofrecer un marco de referencia al hablante oyente para el procesamiento del lenguaje.

La teoría de la invariación acústica tiene importantes implicaciones en la teoría lingüística y en la teoría de la percepción del habla por cuanto ofrece un punto de convergencia entre los principios fonológicos y los modelos de procesamiento auditivo (Blumstein, 1986)<sup>4</sup>.

### *Teoría de la Variabilidad Adaptativa*

Lindblom (1987a, b), por su parte, se plantea el problema de la invariación fonética en los tres dominios de la articulación, la acústica y la percepción con el objeto de elaborar una teoría que pueda explicar cómo realizaciones físicas dispares se interpretan como una unidad por parte del oyente.

---

<sup>3</sup>En relación con la invariación acústica, Stevens (1972a, 1989) propone la naturaleza cuántica de la producción y de la percepción del habla; ciertas articulaciones producen propiedades acústicas estables mientras que otros gestos articulatorios generan resultados acústicos muy sensibles a pequeños cambios de articulación. Los efectos del cambio continuo en la posición articulatoria son, pues, discontinuos en términos del resultado acústico.

<sup>4</sup>En lo que se refiere al ámbito lingüístico, la Teoría de la Invariación Acústica propone un medio de explicar las reglas naturales en fonología; por ejemplo, da cuenta de por qué ciertas reglas de asimilación tienen más probabilidad de aparición que otras. En este sentido, se revisa el proceso de palatalización en las consonantes velares, palatales y dentales del húngaro, proceso que Blumstein considera como un ejemplo de asimilación mientras que Ohala (1986) -comentario a Blumstein, 1986- lo trata como un fenómeno de coarticulación. En cuanto al procesamiento auditivo del lenguaje, se comprueba que las variaciones dependientes del contexto tienen consecuencias perceptivas mínimas si las muestras contienen las propiedades invariables y los parámetros acústicos del habla natural. El comportamiento de [b] y [w] en pruebas de percepción muestra que la distinción fonética entre estos dos sonidos surge de las propiedades acústicas inherentes a cada segmento. Los parámetros acústicos de [b] y [w] son suficientemente distintos por lo que, en el procesamiento de habla, no es necesario contar con los efectos dependientes del contexto.

En cuanto a si la invariación fonética es articulatoria, el autor considera, a partir de la revisión del fenómeno de la reducción formántica ("undershoot"), de los estudios de compensación articulatoria (Lindblom *et al.*, 1987) y del modelo de habla producida a intensidades elevadas ("loud speech") de Schulmann, que las propiedades estables no pueden definirse en términos articulatorios sino que tienen una interpretación más compleja<sup>5</sup>.

En lo que se refiere al carácter acústico de la invariación fonética, Lindblom analiza y contrasta los resultados de los trabajos sobre invariación acústica (Blumstein-Stevens, 1979, Stevens-Blumstein, 1981) y sobre coarticulación (Krull, 1987, 1989). La teoría de la invariación acústica afirma que las propiedades invariables del habla se basan en las características espectrales de la señal en el dominio del tiempo y sin embargo, en las realizaciones orales el fenómeno más relevante es la coarticulación. Por otra parte, según Lindblom, la noción de invariación acústica resulta poco natural ya que los hablantes sólo pueden controlarla a través de sus órganos articulatorios y los oyentes sólo pueden acceder a ella a través de su sistema auditivo (Lindblom, 1987a).

La invariación fonética debe definirse, según Lindblom, en el nivel de la percepción del oyente. Hay indicios importantes relacionados con la interacción hablante-oyente que demuestran que un análisis detallado de la señal no puede, por sí solo, resolver la cuestión de la existencia de elementos invariantes en la lengua (Lindblom, 1986). Por un lado, la comprensión del habla es posible en ausencia parcial de información de la señal debida a la presencia de ruido, reducciones u omisión de indicios. La comunicación presupone en estos casos la aplicación inconsciente del conocimiento lingüístico y conceptual a la interpretación de la señal. Esta estrategia perceptiva que actúa en dirección descendente, desde niveles altos de representación a niveles inferiores, permite a los oyentes restablecer o compensar la información ausente o degradada, al mismo tiempo que explica cómo estímulos físicamente idénticos se perciben de diferente modo según el contexto. Por otro lado, los hablantes son capaces de ajustar sus estilos de pronunciación de acuerdo con las demandas pragmáticas y socio-estilísticas de la situación, elaborando (hiper-articulación) o simplificando (hipo-articulación) su habla en función de las restricciones comunicativas que impone un determinado estilo de habla (Lindblom-Lindgren, 1985). Frente a la teoría de la invariación acústica, Lindblom propone la teoría de la variabilidad adaptativa. El sistema de producción

---

<sup>5</sup>Sin embargo, hay autores (Abbs, 1986; Fujimura, 1986; Harris *et al.*, 1986) que identifican elementos básicos en la producción del habla en el sentido de transformaciones espaciales y temporales. El habla no se genera como una secuencia lineal de unidades discretas sino que es un sistema relacional en el que el indicio más prominente es la existencia de transiciones fonéticas. La invariación articulatoria no se halla, pues, en la posición de los órganos sino en los patrones básicos de sucesión de movimientos.

está regido por dos fuerzas: plasticidad o reorganización orientada al oyente, y economía o simplificación orientada al hablante. Estas dos fuerzas interactúan en el proceso de generación de señales, señales que, según el contexto comunicativo, pueden ser ricas o pobres en información física explícita (cuanto menos informativo es el contexto, más información física deberá contener la señal). El lenguaje es un proceso de adaptación: la variación fonética intra-hablante es genuina y surge como consecuencia del ajuste de la pronunciación del hablante a la situación comunicativa. En este punto, Lindblom coincide con otros autores (Labov, 1972, 1986; Lieberman, 1986; Elman-McClelland, 1986) que reemplazan la búsqueda de elementos invariables en el habla por la aceptación de la existencia de variabilidad no como un atributo que es necesario superar sino como una propiedad intrínseca del lenguaje y que le caracteriza.

Siguiendo a Lindblom, la teoría de la variabilidad adaptativa presenta dos ventajas frente a la teoría de la invariación acústica. En primer lugar, trata la variación fonética como algo intrínseco a todo proceso de habla y en segundo lugar, predice que esta variación ha de estar estructurada de una manera específica, de manera paralela a la organización fonética de los inventarios de sonidos (Lindblom, 1990a; Maddieson, 1984).

En los próximos apartados se relacionarán los resultados del experimento expuestos en anteriores capítulos con la teoría de la invariación acústica y la teoría de la variabilidad adaptativa. Desde el punto de vista de la primera, nos detendremos en la existencia de indicios acústicos que permitan diferenciar entre las categorías analizadas; desde el punto de vista de la segunda, será la interacción de tales indicios con la situación comunicativa el objetivo de la exposición.

#### **4.5.1. INDICIOS ACÚSTICOS**

La búsqueda de indicios acústicos se ha centrado en tres grupos de fenómenos: a) la diferenciación entre hiatos, diptongos y sílabas como entidades globales, b) la discriminación entre vocal, vocal formando parte de un hiato, semiconsonante y consonante, y c) las diferencias entre semivocal y semiconsonante.



#### 4.5.1.1. HIATO vs. DIPTONGO vs. SÍLABA

##### 4.5.1.1.1. DURACIÓN

En el corpus de frases marco, la diferenciación entre un grupo vocálico en hiato o en diptongo en el dominio temporal, se halla en la mayor duración de los hiatos, tanto global como en cada una de sus fases. Es posible además encontrar diferencias debidas a la presencia o ausencia de acento en el caso de los diptongos (diptongos en sílaba tónica vs. diptongos en sílaba átona), y a la posición del acento dentro del grupo en el caso de los hiatos (hiato normal vs. hiato inverso).

En cuanto a la distinción entre hiato, diptongo y sílaba, el hiato palatal es siempre más largo que el diptongo palatal, y la sílaba encabezada por [j] tiende a presentar una duración intermedia entre ambas categorías, aunque se hallan algunas modificaciones de la relación debidas a la interacción del efecto del acento y la vocal adyacente.

En el caso de la comparación de [ue], [ɥe], [we] por un lado y de [ui], [ɥi], [wi] por otro, se da una mayor duración del hiato seguido de la sílaba y del diptongo.

En este sentido, es posible comparar los datos obtenidos en el presente experimento con los ofrecidos por Borzone de Manrique (1979) -v. tabla I-. En el trabajo citado, Borzone de Manrique se plantea como objetivo la descripción y cuantificación de los parámetros acústicos que caracterizan a los diptongos en español. En lo que respecta a la diferencia entre hiatos y diptongos, no halla diferencias en los porcentajes de cambio pero sí en las duraciones de las zonas estables.

	Secuencia vocálica en hiato			Diptongo		
	zona estable	transición	zona estable	zona estable	transición	zona estable
	inicial		final	inicial		final
/io/	88	93	50	35	80	71
/oi/	86	71	92	126	70	63
/ie/	124	63	71	41	63	72
/ei/	28	64	77	50	69	27
/au/	29	73	76	48	55	50
/ua/	76	69	60	43	56	68
/ia/	115	90	58	43	75	73
/ai/	64	77	94	121	71	79
/ue/	73	94	39	39	78	48
/eu/	25	73	91	48	69	36

Tabla I. Duración en milisegundos (ms) de la zona estable inicial, la transición y la zona estable final de un conjunto de secuencias vocálicas en hiato y de un conjunto de diptongos (Borzzone de Manrique, 1979).

Sin embargo, los datos de Borzzone de Manrique están mediatizados por el hecho de que se están comparando pares de palabras en que no sólo se modifica el carácter silábico o asilábico del segmento, sino que también cambia la posición del acento: *ahí* / *hay*. Según esta autora, cuando el acento está situado sobre las vocales /i, u/, en cuyo caso la secuencia vocálica no forma diptongo, las zonas estables de dichas vocales son más largas que las de las vocales abiertas; por el contrario, si el acento cae sobre las vocales abiertas, se establecen otras relaciones. En [i̇a], [ȧi], [i̇o], [u̇e] y [ȧu], las zonas estables de las vocales abiertas son aproximadamente el doble que las del elemento asilábico; en [ȧu], [ėu], [u̇i], [ȯu], aparecen diferencias pero no tan acusadas y finalmente, en [i̇e], [ėi], ambas zonas son similares en duración.

Aquí aislaremos ambas variables (silabicidad/ asilabicidad y acento/ no acento) contrastando conjuntos de datos en que sólo se modifique la variable de la silabicidad; para ello, es necesario tomar en cuenta diptongos acentuados frente a hiatos inversos; así conseguimos combinaciones de semiconsonante, intrínsecamente desprovista de acento, + vocal acentuada por un lado, y vocal no acentuada+ vocal acentuada, por otro lado. No podemos comparar, sin embargo, los hiatos decrecientes con los diptongos decrecientes dado que en

el experimento no se contemplaba tal posibilidad. La tabla II presenta la zona estable inicial, la transición y la zona estable final en secuencias vocálicas en hiato y en diptongos a partir de los datos obtenidos en el análisis del corpus de frases marco.

	Secuencia vocálica en hiato			Diptongo		
	zona estable	transición	zona estable	zona estable	transición	zona estable
	inicial		final	inicial		final
/io/	108.8	21.6	42.77	88.0	20.7	48.8
/oi/						
/ie/	102.87	21.0	49.6	88.6	16.5	42.0
/ei/						
/au/						
/ua/	96.19	22.2	64.0	65.2	15.9	47.0
/ia/	109.6	27.5	66.5	92.6	22.3	46.3
/ai/						
/ue/	128.9	37.4	64.1	76.9	17.5	42.71
/eu/						

Tabla II. Valores medios de duración (ms.) de las secuencias vocálicas en hiato y los diptongos procedentes del corpus de palabras en frase marco.

Las diferencias entre los valores medios pueden explicarse en base a los diferentes diseños experimentales utilizados. Ya se ha mencionado que en los datos de Borzone de Manrique, el efecto de la silabicidad y del acento no están separados; por otro lado, las secuencias proceden de listas de palabras pronunciadas a un ritmo lento mientras que en el presente trabajo las palabras están insertadas en una frase marco que se lee a una velocidad de elocución normal; finalmente, el número de casos difiere de un experimento a otro. En general, los datos de Borzone de Manrique presentan duraciones de transición más largas que los ofrecidos aquí. Si transformamos los valores absolutos en valores relativos, obtenemos las relaciones porcentuales entre las fases de hiatos y diptongos expresadas en la tabla III.

		Borzone de Manrique (1979)					
		Inicial	Transición	Final	Inicial	Transición	Final
HIATO	[io]	38	40	22	63	13	24
	[ie]	48	24	28	59	12	29
	[ia]	44	34	22	54	13	33
	[ua]	37	34	29	53	12	35
	[ue]	35	46	19	56	16	28
DIPT	[io]	19	43	38	56	13	31
	[ie]	23	36	41	60	11	29
	[ia]	23	39	38	58	14	28
	[ua]	26	33	41	51	12	37
	[ue]	24	47	29	56	12	32

Tabla III. Valores en % de las relaciones entre las zonas inicial, de la transición y final de los hiatos y diptongos reinterpretados a partir de los datos de Borzone de Manrique (1979) comparados con los del presente trabajo.

Si bien en los datos del presente trabajo se observa una tendencia por parte de los diptongos a reducir su zona inicial en favor de la zona final, los resultados no son tan evidentes como en el caso de Borzone de Manrique (1979), en que la principal diferencia entre hiatos y diptongos radica en que para los primeros la zona inicial es superior que para los segundos. Sin embargo, la diferencia puede explicarse por la presencia de acento en el conjunto de hiatos frente a su ausencia en el conjunto de diptongos.

Quilis (1981), por su parte, observa que se produce un cambio lento de la transición entre los formantes de las vocales cuando forman parte de un diptongo mientras que un cambio rápido refleja una secuencia heterosilábica. Se basa, sin embargo, en la mera observación de un conjunto de representaciones espectrográficas correspondientes a palabras que incluyen diptongos e hiatos, sin aportar datos que permitan generalizaciones. Por otro lado, ni los datos obtenidos en Borzone de Manrique (1979) ni los del presente trabajo permiten confirmar tal hipótesis: se aprecia una tendencia por parte de los hiatos a presentar un periodo de transición menor al de los diptongos, pero las diferencias no son concluyentes.

#### 4.5.1.1.2. PARÁMETROS FRECUENCIALES

En el nivel de frecuencia, el parámetro determinante en la caracterización de los diptongos es la proporción de cambio de F2, según se ha señalado en diferentes trabajos (Lehiste-Peterson, 1961; Gay, 1968; Burgess, 1969; Borzone de Manrique, 1979, Jha, 1985); también a nivel perceptivo, las posiciones de frecuencia de las zonas estables junto con los porcentajes de cambio de F2 permiten identificar a los diptongos (Gay, 1970). Tal concepción surge a partir de la segmentación de la secuencia en tres partes generalmente denominadas "onglide"- "glide"- "offglide", "onset"- "glide"- "offset", "initial steady state"- "glide"- "final steady state", que posibilita el cálculo de la velocidad de cambio del periodo de transición, especificada como el rango de frecuencia en que un formante se desplaza en un intervalo de tiempo determinado (Lehiste-Peterson, 1961). A partir de aquí, los diptongos se tratan mediante una interpolación, en la mayoría de los casos lineal, entre dos valores de segmentos adyacentes para un mismo articulador (Ren, 1986; Gay, 1970).

Sin embargo, algunos autores (Yang, 1987; Clermont, 1993) han modelizado las trayectorias formánticas de los diptongos mediante ecuaciones que reflejan en mayor medida el comportamiento de los formantes en las secuencias del habla natural.

En el presente análisis, se ha asociado la trayectoria de los formantes en los diptongos a la ecuación polinómica de segundo grado que más se aproxima al conjunto de puntos obtenidos del análisis LPC. Dicho procedimiento se aplica también para distinguir entre las secuencias que están en diptongo de los hiatos.

Si bien es posible hallar diferentes modelos de los diptongos, la distinción entre hiato y diptongo es un problema menos tratado. Según el modelo de truncamiento de Ren (1986), la diferencia entre diptongos e hiatos ha de reflejarse acústicamente. Este autor parte de la idea de que los integrantes de una sílaba están previstos desde el inicio de la misma; según esto, los diptongos tendrán concentrados los componentes en el inicio pero el hiato no, es decir, se dará un truncamiento en el caso del diptongo pero no en el del hiato. El modelo de truncamiento da cuenta de los patrones de F2 en términos de un valor ideal, sin especificación de tiempo pero sí de frecuencia, y de una interpolación de la transición, entendida como la conexión de dos valores ideales fonológicamente adyacentes, con una

proporción de cambio especificada. Según Ren (1986), añadir un componente a la sílaba no se realiza comprimiendo los componentes originales de la misma, sino truncando la parte inicial de la sílaba para permitir una conexión dinámica entre el nuevo componente y el primer componente de la sílaba inicial, mientras que el resto de la sílaba permanece intacta.

Borzone de Manrique (1979), por el contrario, no halla diferencias relevantes en los valores de proporción de cambio formántico entre las secuencias vocálicas que forman diptongo y las que no: sólo la duración de las partes estables varía. Sin embargo, el procedimiento de análisis de Borzone de Manrique, basado en la obtención de los valores de F1, F2 y F3 en los puntos más cercanos a los valores ideales elimina toda la información acerca de la dinámica de la trayectoria de los formantes. En el presente trabajo, se obtienen diferencias importantes entre diptongos e hiatos en la curvatura de F2: la trayectoria de F2 en los diptongos presenta una curva más pronunciada. Por otro lado, cuando los segmentos vocálicos pertenecen a diferentes sílabas se da una influencia mutua menor, mientras que cuando es un grupo tautosilábico, se da un reajuste de los valores.

Así como en el estudio de van Bergem (1993b) sobre la dinámica de los formantes en las vocales se observa que el cambio de las frecuencias de los formantes en las zonas estables de las vocales debido a un cambio de estilo de habla coincide con una curva más reducida de las trayectorias, la distinción entre hiato y diptongo en el presente trabajo se ha relacionado con una curva más pronunciada para los diptongos, además de que el cambio de situación comunicativa incide en la reducción del grado de curvatura de hiatos y diptongos.

Por otro lado, dado que en los diptongos se observa en general una trayectoria formántica ininterrumpida desde la consonante inicial a la final, frente a dos zonas relativamente estables, correspondientes a las vocales, en los hiatos, se espera un menor ajuste de la curva de la ecuación polinómica en los hiatos que en los diptongos. La medida de ajuste a la curva viene dada por el índice de correlación ( $r$ ). Excepto en el caso del primer formante de las secuencias palatales, en que un análisis ANOVA no halla diferencias ( $p=.61212$ ) en el resto de comparaciones el mismo tipo de análisis señala diferencias a un nivel de significación del 99% entre el ajuste a la curva por parte de los puntos de análisis de los diptongos frente a los hiatos; tal situación se repite para las secuencias procedentes de la tarea del mapa, esta vez incluyendo también la frecuencia de F1 de los grupos palatales. La tabla IV expone los valores medios del índice de correlación de la curva respecto al conjunto de puntos de análisis de F1 y F2 de los hiatos y diptongos palatales en el corpus de frases marco y en el de la tarea del mapa; la tabla V muestra la misma información relativa a los grupos velares.

	LISTAS				TAREA DEL MAPA			
	F1		F2		F1		F2	
	HIATO	DIPT	HIATO	DIPT	HIATO	DIPT	HIATO	DIPT
ia	.95	.95	.96	.97	.94	.95	.94	.95
ie	.88	.89	.92	.94	.86	.93	.84	.95
io	.87	.88	.93	.96	.80	.92	.92	.94
	.87	.91	.94	.96	.87	.93	.91	.95

Tabla IV. Valores medios del índice de correlación de la curva al conjunto de puntos de análisis de F1 y F2 en secuencias vocálicas en hiato y en diptongos palatales en el corpus de frases marco y en el de la tarea del mapa.

	LISTAS				TAREA DEL MAPA			
	F1		F2		F1		F2	
	HIATO	DIPT	HIATO	DIPT	HIATO	DIPT	HIATO	DIPT
ua	.89	.95	.86	.93	.95	.98	.90	.98
ue	.88	.92	.88	.92	.97	.97	.98	.96
uo	.92	.93	.82	.88	.96	.96	.90	.95
	.90	.93	.86	.91	.96	.97	.94	.96

Tabla V. Valores medios del índice de correlación de la curva al conjunto de puntos de análisis de F1 y F2 en secuencias vocálicas en hiato y en diptongos velares en el corpus de frases marco y en el de la tarea del mapa.

La modelización por medio de una curva se revela adecuada en la caracterización de los diptongos, además de que ofrece un índice de discriminación frente a las secuencias heterosilábicas: hiatos y diptongos presentan una forma de curva diferente, y el grado de ajuste al modelo por parte de los primeros es menor.

#### 4.5.1.2. VOCAL *vs.* VOCAL EN HIATO *vs.* SEMICONSONANTE *vs.* CONSONANTE

##### 4.5.1.2.1. DURACIÓN

El comportamiento temporal de los segmentos considerados en el experimento de palabras en frase marco (vocal en entorno consonántico, vocal en contacto con otra vocal, semiconsonante y consonante) permite establecer diferencias entre los mismos de tal modo que, independientemente de la vocal adyacente y del entorno acentual, una vocal [i, u] en contacto con otra vocal presenta siempre mayor duración que la misma vocal en entorno consonántico y que la correspondiente semiconsonante. En cuanto a la consonante, en la serie palatal presenta mayor duración que la semiconsonante, mientras que en la serie velar muestra una duración inferior. No hay, sin embargo, diferencias duracionales entre la semiconsonante, la vocal palatal y la consonante palatal, por un lado, ni entre la semiconsonante y la vocal velar, por otro lado. Únicamente se hallan diferencias entre la semiconsonante y la vocal velar en ambos estilos.

Sin embargo, Williams (1982) identifica un rasgo fonético "semi-" frente a las vocales, caracterizado acústicamente por la aparición de unas bandas formánticas de duración menor. La tabla VI presenta los valores medios de los elementos asilábicos y de los núcleos de los diptongos a partir de Williams (1982) y el presente trabajo; la tabla VII presenta, por su parte, los valores medios de las vocales [i], [u] en contexto acentuado en ambos estudios. Williams (1982) interpreta los resultados de modo que la comparación entre las vocales en un entorno consonántico y los elementos asilábicos del diptongo resultan en una tendencia a una mayor duración por parte de éstos últimos. Sin embargo, la comparación entre los elementos asilábicos y el núcleo del diptongo indica una mayor duración de la vocal. El problema radica en la falta de información acerca del número de casos analizados, los criterios de segmentación y la significación estadística de los valores. Por otro lado, la simple observación de los datos no apoya en todos los casos las afirmaciones antes expuestas: en [aʊ], el segmento asilábico y el núcleo presentan idéntica duración, y en [eʊ], [aɪ#], [eɪ#], el núcleo es más corto que la semivocal.



Williams (1982) (* *)				
	Segmento	Núcleo	Segmento	Núcleo
	asilábico		asilábico	
[aɪ]#	93	40	(*)	
[eɪ]#	80	77	(*)	
[oɪ]#	62	100	(*)	
[oɪ]	46	93	39.3	85.6
[eɪ]	77	67	42.8	59.2
[aɪ]	46	65	46.2	75.5
[iɛ]	40	62	88.6	42.1
[iɔ]	30	62	88.0	48.8
[aʊ]	62	62	50.3	79.8
[eʊ]	77	74	58.1	88.2
[ʊe]	12	65	76.9	42.1
[ʊa]	12	77	65.2	47.0
[ui]	67	94	68.9	35.8

Tabla VI. Valores medios de los elementos asilábicos y de los núcleos en los diptongos en Williams (1982) y en el presente trabajo.

(\*) Los diptongos en posición final no han sido considerados en el presente experimento.

(\*\*) Los datos están presentados en segundos en el trabajo original; aquí se reinterpretan en milisegundos..

Williams (1982)		
[i] +ac	30	84.25
[u] +ac	73.8	93.68

Tabla VII. Valores medios de duración de las vocales [i], [u] en entorno consonántico en Williams (1982), a partir del promedio de las diferentes estructuras silábicas, y en el presente análisis.

También Monroy Casas (1980) caracteriza a las semiconsonantes por su duración breve y por una fricción palatal apenas perceptible, aunque se basa en la observación de una serie de espectrogramas, que por otra parte, tampoco parecen apoyar sus afirmaciones; nuevamente,

el problema radica en saber cuáles son los criterios de segmentación adoptados para decidir las fronteras entre el elemento asilábico y el núcleo.

Los resultados del presente análisis no permiten relacionar de forma directa la mayor duración del segmento con la noción de núcleo silábico, dado que la tendencia es que el elemento inicial, sea silábico o asilábico, presente mayor duración. Los valores de duración de las vocales en entorno consonántico, por otra parte, se asemejan a los de los elementos iniciales.

#### 4.5.1.2.2. PARÁMETROS FRECUENCIALES

Borzzone de Manrique (1976) se plantea la descripción acústica de las variantes inicial absoluta, inicial tras consonante y final de /i, u/ en los diptongos [i̯e], [u̯e] del español de Argentina. Los formantes de la variante en C+[i̯e] coinciden con los de C+[i] mientras que F1 y F2 de [u] en C+[u̯e] se sitúan en zonas de frecuencia superiores a los de C+[u]. Tales resultados coinciden con los del presente análisis en lo que se refiere a la variante velar, pero también se han señalado modificaciones frecuenciales en la semiconsonante palatal. El elemento semiconsonántico presenta una mayor tendencia a la dispersión y un desplazamiento de los dos primeros formantes con respecto a la vocal en hiato y a la vocal en contexto consonántico. El desplazamiento formántico se traduce en una elevación de F1 y un descenso de F2 en la semiconsonante palatal y en una elevación de F1 y de F2 en la semiconsonante velar.

En cuanto a la variante inicial absoluta palatal, en Borzzone de Manrique (1976) se caracteriza por el descenso de F1, el aumento de F2 y de F3, la caída de intensidad con respecto al siguiente elemento vocálico, y la presencia de elementos de ruido en el inicio; la variante velar, por su parte, presenta un descenso de F1 y de F2, junto con una menor intensidad relativa a la vocal adyacente [e]. Esta variante es claramente distinta al resto y debe describirse como una semi-vocal [i̯], [u̯] en discrepancia con Navarro Tomás (1918) que la considera como una consonante. También se sitúa en discrepancia con los resultados del presente experimento, dado que se ha considerado tal variante como alófono consonántico en base a las diferentes realizaciones fonéticas halladas: en posición inicial absoluta, aparecen variantes palatales oclusivas, fricativas -con y sin prenasalización- y aproximantes,

ésta última muy cercana acústicamente (en el dominio temporal y de la frecuencia) a la variante semiconsonántica; en cuanto a la variante velar, se halla una alternancia [w], [ɣ̞], [g̞].

En su análisis del español de Lima, Williams (1982) caracteriza a [j] y a [w] como consonantes fricativas<sup>6</sup>: [j] se caracteriza acústicamente por la presencia de fricción y de sonoridad; en cuanto a [w], no hay verdadera fricción, pero una serie de indicios la diferencian de la vocal [u] -formante de sonoridad, transición de F1 más rápida, inicio de F3 posterior al resto de formantes-. También Quilis (1981, 1993) clasifica a [j] como fricativa, junto a [β ð ɣ]: un descenso de F1 y de energía con respecto a la vocal contigua, así como unas transiciones lentas la caracterizan, según este autor. La clasificación de Quilis (1981) y de Williams (1982) de tales segmentos como fricativos no se basa en la presencia de fricción sino principalmente en las posiciones de los formantes alcanzadas durante la producción del sonido.

Sin embargo, en el presente trabajo se han podido aislar variantes palatales fricativas, con presencia de ruido por encima de los 3000 Hz. y variantes palatales aproximantes, con estructura formántica definida. En cuanto a la variante velar, se realiza como una aproximante, distinta de [ɣ] por la presencia de labialización, que se manifiesta acústicamente en forma de dos formantes muy próximos en la zona baja del espectro. La consonante [w] muestra un comportamiento diferenciado frente a la semiconsonante [ɰ] y a la vocal [u]: las frecuencias de F1 y de F2 de la consonante se sitúan en zonas inferiores, lo cual señala una realización cerrada y labiovelar.

Por último, el efecto de coarticulación sobre la vocal adyacente es más importante en el caso de la consonante que en el resto de segmentos en la serie velar, aunque no en la serie palatal. Las vocales contiguas a [w] presentan un grado mayor de velarización que cuando aparecen junto a [ɰ] o a [u].

---

<sup>6</sup> Hay que tener en cuenta en este caso las diferencias de realización fonética entre el español peninsular y el citado (Zamora, 1960; Canfield, 1981).

#### 4.5.1.3. SEMIVOCAL vs. SEMICONSONANTE

En cuanto a la duración del diptongo en función de si es creciente o decreciente, Monroy Casas (1980), a partir del análisis de una serie de palabras pronunciadas por dos informantes españoles, observa que los diptongos crecientes presentan menor duración que los decrecientes. Los datos del presente trabajo no corroboran de forma general tal afirmación: en el corpus de listas, los diptongos crecientes palatales son más largos que los decrecientes palatales, mientras que los decrecientes velares muestran una mayor duración que los correspondientes crecientes; en el corpus del mapa, no se dan diferencias en los grupos palatales en función de la posición del elemento asilábico, pero en los grupos velares la duración de los diptongos decrecientes es superior a la de los crecientes.

Según los resultados del presente experimento la semivocal se diferencia de la semiconsonante en la duración y en las frecuencias de los dos primeros formantes: la semiconsonante es más larga que la semivocal, y presenta un valor de F1 más bajo y un valor de F2 más alto. La descripción frecuencial coincide con la de Borzone de Manrique (1976) y Williams (1982). Borzone de Manrique (1976) observa que mientras que los formantes de [i] en el contexto C+[i̯e] coinciden con los de [i] en C+[i], la variante en el contexto C+[e̯i] sufre modificaciones de frecuencia en el sentido de una elevación de la frecuencia de F1 y un descenso de F2 y F3. También se da desplazamiento de frecuencias en la comparación entre la variante C+[ɥe] y C+[eɥ]. La semiconsonante y la vocal tienen el mismo grado de abertura mientras que la semivocal es más abierta.

Williams (1982), por su parte, presenta una descripción frecuencial similar pero en cuanto a la temporal señala una mayor duración de la semivocal frente a la semiconsonante.

A la descripción puramente acústica, se suma el efecto que el segmento ejerce sobre los elementos adyacentes. El efecto de coarticulación de semiconsonante y semivocal no es el mismo: la semiconsonante ejerce un efecto muy fuerte de palatalización sobre la vocal siguiente, mientras que la semivocal no muestra dicha influencia.

En cuanto a las combinaciones [iu]-[ui], es posible diferenciar entre los diptongos, hiatos y sílabas en el dominio temporal y de la frecuencia. Lehiste-Peterson (1961) observan que en las secuencias [ui]/ [iu] del finlandés la proporción de cambio formántico determina el contraste vocal-vocal / semivocal-vocal; también atribuyen a la diferencia en las transiciones y a las duraciones de las posiciones estables la distinción entre los diptongos [ui]-[iu] del estoniano frente a las secuencias inglesas [wi]-[ju].

Sin embargo, en el momento de discernir cuál de los elementos es asilábico y cuál núcleo en los diptongos formados por dos vocales cerradas, los parámetros acústicos no aportan argumentos decisivos para decidir si [i] en [iu] es semiconsonante o núcleo, y si [u] en [iu] es núcleo o semivocal. En cualquier caso, aparecen indicios tanto en el corpus de frases marco como en el de la tarea del mapa que favorecen la consideración de [u] como núcleo en [iu] y en [ui]. Tales argumentos se refieren a la duración de [u] equivalente a la de los diptongos decrecientes velares y a los valores formánticos de [u] en [iu], similares a los de una vocal y no a los de una semivocal. En este sentido, se coincide con las descripciones clásicas del español en que no se admite la existencia del diptongo [iu] (Alarcos, 1965; Canellada-Kuhlmann, 1987; Martínez de Sousa, 1974); en cuanto a [ui], sin embargo, el acuerdo no es general. Canellada-Kuhlmann (1987) aportan como argumento en favor de [ui] la pronunciación enfática en que se alarga [u] y no [i]: "muy bien, pero que MUY bien ['mu:i'β̞ien] y no [mui:β̞ien]". Las reducciones del habla vulgar aportan argumentos adicionales: "muy" se manifiesta como ['mu] y no como [mi].

En cuanto a las combinaciones en hiato, según los tratados de la lengua española, las secuencias ['iu], ['ui] no están permitidas en la lengua, dado que en la unión de dos vocales altas el acento siempre recae en la segunda (Navarro Tomás, 1918; R.A.E., 1973). Dicha afirmación queda corroborada en el presente experimento en que todos los informantes han pronunciado tales secuencias como [i'u], [u'i].

La diferencia entre semivocal y semiconsonante viene motivada por el efecto de la posición en la sílaba y por las tendencias coarticulatorias de dirección progresiva que muestra la lengua española. De hecho, diptongos crecientes y decrecientes son imágenes especulares y las diferencias entre semivocal y semiconsonante pueden explicarse por el cambio de dirección de las trayectorias sin recurrir al establecimiento de propiedades intrínsecas de los sonidos. Siguiendo la intuición de Cressey (1975), *"there is a distinction between an allophonic difference in a particular language (e.g. [b]-[β] alternations in Spanish) and the*

*difference between semivowels and semiconsonantes which is the "universal" result of the contact with other sounds. That is, one can imagine a language without the allophonic conditioning suffered by Spanish /b/, /d/, /g/, or even with the reverse effects, but it is inconceivable that a language might have sequences such as CSvV or VScV*"<sup>7</sup>. Tal intuición se materializa como la imposibilidad de alcanzar el valor ideal: la trayectoria de los formantes se trunca antes de alcanzar el punto final. La misma idea viene apoyada por los estudios sobre la influencia de la velocidad de elocución en los diptongos decrecientes del inglés americano: *"The articulatory gesture begins at a fixed position in the vowel space. This position is held for a particular length of time, the duration of which is governed by the speaker's rate of production and/ or the presence or absence of a following voiced consonant. The articulators then begin to move toward a final target position at a given speed. If speaker rate is slow, the target is reached and the gesture is completed: if speaker rate is fast, the movement, while on course, is cut off before reaching the final target"* (Gay, 1970<sup>8</sup>)

#### 4.5.2. INTERACCIÓN CON LA SITUACIÓN

Una de las principales características de un estilo de habla conversacional es la reducción temporal en comparación con las secuencias procedentes de otros estilos como la lectura de listas de palabras o de frases (cf. Eskénazi, 1993).

En el experimento presentado, todas las secuencias muestran una duración menor cuando se extraen del corpus de frases marco que cuando se obtienen del corpus de la tarea del mapa; en general, los hiatos reducen su duración un 16%, los diptongos el 17% y las sílabas el 26%. Sin embargo, las secuencias integradas por un elemento velar presentan mayor reducción que las formadas por un elemento palatal: se da una reducción del 12% en los hiatos palatales y del 23% en los velares, del 18% en los diptongos palatales frente al 23% en los velares; del 32% en las sílabas encabezadas por [j] y del 35% de las sílabas [w]V. En

---

<sup>7</sup>nota 2, p ág. 41

<sup>8</sup>pág.. 1573

Borzzone de Manrique (1979) el cambio de una condición a otra -palabras aisladas, frase marco- implica una reducción de la duración total entre el 25 y el 50%; Jha (1985), por su parte, señala que un cambio en la velocidad de elocución conlleva una reducción entre el 38 y el 64% de la duración total.

Sin embargo, pese a tal reducción en duración, se mantiene la mayoría de relaciones frecuenciales y temporales entre los segmentos. En ambos conjuntos de datos, los grupos vocálicos mantienen una relación de alargamiento del hiato con respecto al diptongo.

En trabajos anteriores, la única variable analizada en la modificación de los diptongos ha sido la velocidad de elocución. Gay (1968) señala que la frecuencia inicial y el porcentaje de cambio de la transición de F2 en la producción de los diptongos del inglés americano es constante a través de las diferencias de velocidad de elocución. Borzzone de Manrique (1979) y Jha (1985) aportan datos experimentales adicionales que parecen confirmar esta hipótesis: tanto en los diptongos del español como en los del maithili, el porcentaje de cambio se mantiene a pesar de los cambios de duración.

Sin embargo, Toledo-Antoñanzas Barroso (1987) obtienen nuevos datos para el español de Argentina que rebaten el modelo de Gay (1968): señalan diferencias importantes en el porcentaje de cambio de F2 debidas a los cambios de velocidad de elocución. También la frecuencia de F2 inicial muestra variaciones relevantes en los diptongos [ja eu ue oi io]. En la comparación, sin embargo, entre el grupo de trabajos citados interviene un factor dependiente del diseño experimental: se encuentran dos concepciones diferentes de lo que se considera cambio de velocidad de elocución. Si bien Gay (1968) y Toledo-Antoñanzas Barroso (1987) obtienen los datos a partir de la lectura de las palabras insertadas en frase marco, a tres velocidades de elocución -lenta, moderada y rápida-, la tónica general es identificar la variable "velocidad de elocución" con el contexto en que aparece la secuencia que se va a analizar. Borzzone de Manrique (1979) plantea dos condiciones experimentales relacionadas con la velocidad de elocución: las palabras aisladas se asocian a una velocidad de elocución lenta y las palabras insertadas en frases marco a una velocidad rápida.<sup>9</sup> Jha (1985), por su parte, analiza los diptongos [əi], [əu] del maithili pronunciados en tres condiciones experimentales: secuencias aisladas, insertadas en palabras y en frases marco.

---

<sup>9</sup> Aquí no coincidimos con tal asociación al considerar que la variable contexto -palabras aisladas o insertadas en frase- no debe identificarse con la variable velocidad de elocución -rápida, moderada, lenta- (cf. capítulo 4.1.).

El autor, en la misma línea que Borzone de Manrique (1979), relaciona dichas condiciones con una velocidad de elocución lenta, moderada y rápida, respectivamente.

En general, cuanto más rápida es la velocidad de elocución, más se aleja el valor de frecuencia formántica del ideal: en las palabras y en las frases marco, el segundo elemento no alcanza la posición de frecuencia de una vocal. Sin embargo, en esta comparación se presupone que el elemento que forma parte del diptongo es una vocal (Borzone de Manrique, 1979; Jha, 1985).

En el presente trabajo no se considera el cambio de velocidad de elocución ni el contexto (aislado, frase, párrafo) en que aparece la secuencia, sino cómo el cambio de situación comunicativa influye en la realización de las secuencias. Por tanto, junto a la reducción de duración, se ponen de manifiesto otros fenómenos, como la pérdida de distinciones acústicas en el habla conversacional y la presencia de reducciones fonéticas.

En cuanto a la atenuación de indicios acústicos, las relaciones acento-duración no son siempre evidentes en el habla conversacional, a la vez que se neutralizan las distinciones en los grupos vocálicos en función de la vocal adyacente. En otros trabajos y en torno a otros segmentos fonéticos se ha puesto de manifiesto la misma tendencia: discriminación pobre de las vocales del español obtenidas de un monólogo (Harmegnies-Poch, 1991, 1992), aumento del grado de coarticulación en las consonantes (Krull, 1987, 1989; Machuca, 1991; de la Mota, 1991; Aguilar-Andreu, 1991). Por otro lado, se advierten modificaciones de las relaciones en las curvaturas de F1 y F2 de los grupos vocálicos en hiato frente a los correspondientes diptongos. Nos centraremos en el segundo formante, cuyo desplazamiento se ha tomado como índice de la reducción formántica (Lindblom, 1963; van Bergem, 1993). Aparece una pendiente de F2 menos pronunciada en los grupos vocálicos palatales en el corpus de lectura de listas frente a una pendiente más pronunciada de los grupos vocálicos velares. La figura 1 presenta en forma esquemática el proceso. En el corpus de la tarea del mapa, los valores ideales no se alcanzan, lo cual modifica la forma de la curva; las trayectorias en los diptongos son más planas en las secuencias de habla conversacional, confirmando la idea de la existencia de reducción formántica en los estilos de habla relajados, así como la relación inversa entre duración y frecuencia formántica: tal como decrece la duración, las posiciones de los formantes están más próximas en el espacio vocálico.



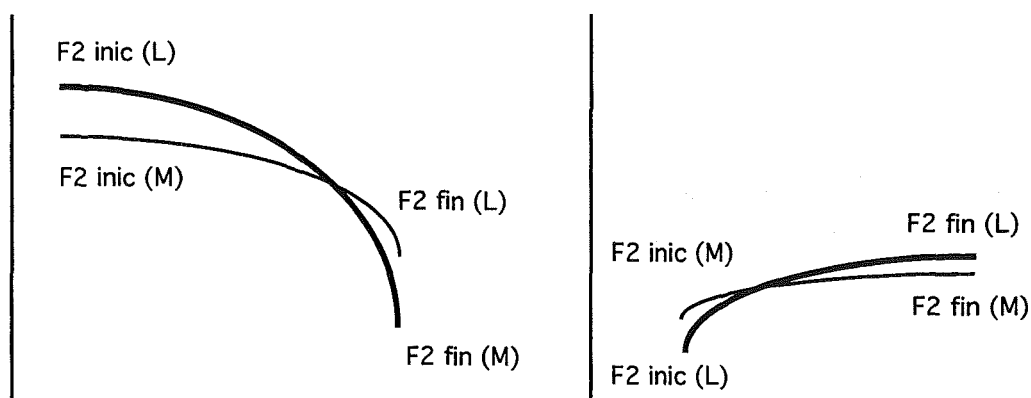


Figura 1. Esquemización de las trayectorias formánticas de F2 en los diptongos palatales (izquierda) y velares (derecha) en el corpus de listas (trazo grueso) y en el de la tarea del mapa (trazo delgado).

En cuanto a la presencia de reducciones fonéticas, se ha puesto de manifiesto un eje de debilitamiento o reducción en los grupos vocálicos, que va desde la presencia de un hiato a una vocal, pasando por el estadio intermedio del diptongo. La existencia de un continuo de reducción también se ha revelado en el comportamiento de las consonantes (Kohler, 1989; Aguilar *et al*, 1993): en habla informal, cualquier consonante puede verse sometida a un proceso de reducción que además presenta diferentes estadios -por ejemplo, las oclusivas sordas del español pueden manifestarse como oclusivas sonoras, aproximantes o elidirse. Por otra parte, las reducciones en los grupos vocálicos se han puesto de manifiesto también en el nivel interlexémico, tanto en las combinaciones en las que interviene una vocal alta como en las combinaciones de dos vocales no altas (Aguilar, 1991; Aguilar-Machuca, 1993a, 1993b).

Desde un enfoque más sociolingüístico, Burgess (1969) observa las realizaciones de /i/, /eI/, /aI/, /au/, /ou/, /u/ en dos registros de habla en el inglés australiano, uno caracterizado por una variedad cultivada y otro por una popular. El análisis basado en espectrogramas señala diferencias de realización entre ambos grupos de hablantes. Ingram (1989), por su parte, también cita como característica de los adolescentes menos cultivados la presencia de monoptongación de los diptongos del inglés australiano.

### 4.5.3. CONCLUSIONES

Los resultados del experimento en un corpus de lectura de palabras junto con los resultados procedentes de un habla en la que la atención sobre el propio lenguaje se pierde, ofrecen información acerca de las propiedades acústicas de determinados segmentos así como en torno a las modificaciones que sufren tales propiedades por efecto de un cambio en la situación comunicativa. Es posible hallar unos invariantes acústicos asociados a las categorías de vocal, semiconsonante, semivocal y consonante, además de una serie de procesos de debilitamiento de indicios acústicos y de reducción fonética, que ponen de manifiesto las tendencias articulatorias hacia un relajamiento en habla menos cuidada.

Se argumenta por tanto en favor de la teoría de la variabilidad adaptativa formulada por Lindblom (1986, 1987 a, 1990 b). La producción del habla depende del principio del máximo contraste perceptivo con el mínimo esfuerzo articulatorio. En el corpus de la tarea del mapa, la tendencia a la reducción fonética es muy acusada debido a que se trata de un diálogo: la presencia de un interlocutor permite al informante apoyarse en el conocimiento lingüístico y pragmático de modo que su pronunciación se relaja. La estrategia perceptiva humana, siguiendo a Lindblom, actúa desde niveles altos de representación a niveles inferiores, lo cual permite al oyente restablecer la información acústica inexistente -por ejemplo, las vocales elididas- o degradada -transmisión comunicativa en ambientes ruidosos-. Por el contrario, en la lectura de frases marco no se presupone una audiencia y el informante ajusta su estilo de pronunciación a la nueva situación. El resultado es que en la conversación se advierte una tendencia a la hipo-articulación o simplificación articulatoria en beneficio del emisor, mientras que en lectura domina la hiper-articulación o elaboración de los gestos implicados en el proceso de producción de acuerdo con la demanda pragmática. En otros términos, se aportan así argumentos en favor de la distinción de dos modos de producción: la lectura y el habla, presentada en capítulos anteriores. Por otra parte, de acuerdo con Lindblom (1990b), la cuestión de la invariación fonética se sustituye por otra tarea: "*explicating the notion of sufficient discriminability*". Lo que el hablante controla no es que las unidades lingüísticas se actualicen en términos de invariantes físicas, sino que los atributos de su señal posean suficiente contraste, es decir que su poder de discriminación baste para el acceso léxico. Así, las propiedades puestas de manifiesto para vocales, semiconsonantes y consonantes aproximantes se configuran como los contrastes mínimos y suficientes para su identificación.

## 5. CONCLUSIONES

---

Las conclusiones globales del presente trabajo se estructuran en dos niveles: la discusión de las variantes vocálicas y consonánticas que han aparecido en el análisis acústico, con la consiguiente revisión del inventario de sonidos del español, y la clasificación de dichas variantes con respecto a las entidades fonémicas de la lengua.

La asignación de los sonidos [i], [i̠], [j], [J], [u], [u̠], [w] a unos fonemas u otros se basa en un examen de sus características fonéticas y de sus propiedades de distribución en la lengua. Los argumentos proceden en su mayoría de la fonémica clásica -mínimo contraste, similitud fonética, distribución- aunque incorporamos también algunas propiedades relacionadas con la estructura silábica, enfatizadas desde el generativismo, además de algunos datos externos, argumentación habitual en fonología natural y de laboratorio. El objetivo es dilucidar el carácter fonético y el comportamiento fonológico de unos segmentos polémicos en español, con el fin de obtener generalizaciones aplicables en dominios de la tecnología del habla como sistemas de transcripción automática (Ríos, 1993) o sistemas de conversión texto a habla y de reconocimiento. El establecimiento de un repertorio de sonidos más cercano al del habla utilizada en situaciones cotidianas, así como la organización de los mismos en clases mayores o fonemas puede introducir mejoras en los ámbitos mencionados.

## 5.1. VARIANTES FONÉTICAS

### 5.1.1. VARIANTES VOCÁLICAS

A continuación se presentan las variantes vocálicas que se han puesto de manifiesto en el análisis de las secuencias en dos conjuntos de datos- lectura de palabras insertadas en frases marco, pronunciaciones obtenidas en un conjunto de diálogos-<sup>1</sup>. Junto a la descripción fonética se ofrece una breve caracterización acústica y la distribución que se desprende a partir del análisis del corpus.

Con el fin de disponer de valores de referencia correspondientes a las realizaciones que se van a describir, la tabla I expone los valores medios de frecuencia obtenidos en el análisis de las secuencias extraídas de las frases marco y la tabla II los correspondientes a las muestras de la tarea del mapa.

Para facilitar además la interpretación de los datos incluimos los resultados de las pruebas estadísticas en la tabla III.

	[i]	[i]V	[i̥]V	V[i]	[u]	[u]V	[u̥]V	V[u]
F1	325	335	337	389	337	338	354	378
F2	2219	2150	2113	2000	803	837	828	989

Tabla I. Valores medios de F1 y de F2 de las variantes vocálicas consideradas en el análisis de secuencias procedentes de un corpus de palabras en frases marco.

<sup>1</sup> Seguimos utilizando los símbolos propuestos al inicio del trabajo; la discusión de los mismos en función de los datos obtenidos en el análisis ocupará la sección 5.1.3.

	[i]	[i]V	[ị]V	V[ị]	[u]	[u]V	[ụ]V	V[ụ]
F1	348	349	360	397	356	361	374	406
F2	2104	2117	2019	1904	891	892	872	1041

Tabla II. Valores medios de F1 y de F2 de las variantes vocálicas consideradas en el análisis de secuencias procedentes del corpus de la tarea del mapa.

	LISTAS		TAREA DEL MAPA	
	[i] vs. [i]V vs. [ị]V vs. V[ị]	[u] vs. [u]V vs. [ụ]V vs. V[ụ]	[i] vs. [i]V vs. [ị]V vs. V[ị]	[u] vs. [u]V vs. [ụ]V vs. V[ụ]
F1	***	***	***	***
	* VH vs. SV	VH vs. SC	* VH vs. SV	*VH vs. SC
	VH vs. SC	VH vs. V	VH vs. SC	VH vs. V
	VH vs. V	* VH vs. SV	VH vs. V	* VH vs. SV
	* SC vs. V	SC vs. V	SC vs. V	*SC vs. V
	SC vs. SV	* SC vs. SV	* SC vs. SV	* SC vs. SV
	* SV vs. V	* SV vs. V	*SV vs. V	* SV vs. V
F2	***	***	***	***
	* VH vs. SC	VH vs. SC	VH vs. SC	VH vs. SC
	* VH vs. SV	VH vs. V	* VH vs. SV	* VH vs. SV
	* VH vs. V	* VH vs. SV	VH vs. V	VH vs. V
	* SC vs. SV	* SC vs. V	* SC vs. SV	* SC vs. SV
	* SC vs. V	* SC vs. SV	SC vs. V	SC vs. V
	* SV vs. V	* SV vs. V	* SV vs. V	* SV vs. V

Tabla III. Diferencias estadísticamente significativas (\*\*\*)  $p < .0005$  ---  $p > .0005$ ) en los parámetros de F1 y de F2 de vocal (V) vs. vocal formando parte de un hiato (VH) vs. semiconsonante (SC) vs. semivocal (SV) en el corpus de palabras en frases marco y en el de la tarea del mapa.

---

[i]	Vocal anterior cerrada
Características acústicas	Presencia de dos formantes muy separados, el primero en zonas bajas del espectro y el segundo en zonas altas de frecuencia.
Distribución	Aparece como vocal silábica en cualquier contexto.

---

---

[i:]	Vocal anterior cerrada semilarga
Características acústicas	Zonas de frecuencia similares a las de [i], pero presencia de una mayor duración.
Distribución	Aparece como vocal formando parte de un grupo vocálico en hiato.

---

---

[ɨ]	Vocal anterior semicerrada
Características acústicas	Valores de F1 superiores y valores de F2 inferiores a los de [i]: acústicamente similar a [j] en duración y en frecuencia de F1 y de F2.
Distribución	Aparece como elemento inicial en un diptongo.

---

---

[ɨ]	Vocal semianterior semicerrada
Características acústicas	Valores de F1 superiores a los de [i] y a los de [i̞] prevocálica; valores de F2 inferiores a los de [i] y a los de [i̞] prevocálica.
Distribución	Aparece como elemento final en un diptongo.

---

---

[u]	Vocal posterior cerrada
Características acústicas	Presencia de dos formantes muy próximos situados en la zona inferior del espectro
Distribución	Aparece como vocal silábica en cualquier contexto.

---

---

[u:]	Vocal posterior cerrada semilarga
Características acústicas	Zonas de frecuencia similares a las de [u], pero presencia de una mayor duración.
Distribución	Aparece como vocal formando parte de un grupo vocálico en hiato.

---

---

[ɯ]	Vocal posterior semicerrada
Características acústicas	Valores de F1 superiores a los de [u] y valores de F2 inferiores a los de [u].
Distribución	Aparece como elemento inicial en un diptongo.

---

---

[u̠]	Vocal posterior semicerrada
Características acústicas	Valores de F1 y de F2 superiores a los de [u] prevocálica y a los de [u].
Distribución	Aparece como elemento final en un diptongo.

---

### 5.1.2. VARIANTES CONSONÁNTICAS

El análisis acústico de las secuencias del corpus pone de manifiesto la existencia de una serie de realizaciones consonánticas diferenciadas. Si bien dicho análisis no tenía como principal objetivo la caracterización acústica de tales variantes fonéticas, en el presente apartado se establecerán las propiedades acústicas y la distribución de las mismas.

A continuación se enumeran las variantes consonánticas aisladas, junto con su caracterización acústica y su distribución en el conjunto de datos observado. La enumeración no pretende ser exhaustiva pero aporta información relevante sobre el inventario fonético del español. Los porcentajes de frecuencia de aparición y los condicionamientos fonéticos se han detallado en anteriores capítulos (cf. 4.4).

---

[d̪]	Africada palatal sonora
Características acústicas	Periodo de oclusión con sonoridad seguido de un periodo de fricción.
Distribución	Aparece en posición inicial absoluta como manifestación de un reforzamiento por parte de algunos hablantes en una pronunciación cuidada.

---



---

[ʃ]	Oclusiva palatal sonora
Características acústicas	Formante de sonoridad seguido de una explosión.
Distribución	Aparece en posición inicial absoluta en la lectura de frases.

---

---

[ʒ]	Fricativa palatal sonora
Características acústicas	Zona de fricción que se inicia en torno a los 3000 Hz. acompañada de sonoridad
Distribución	Aparece en posición inicial absoluta en la lectura de frases. En el habla conversacional se manifiesta como preferencia por parte de algunos informantes frente a la variante aproximante

---

---

[j]	Aproximante palatal sonora
Características acústicas	Estructura formántica definida y amplitud similar, aunque inferior, a la de la vocal adyacente
Distribución	Aparece con carácter general en posición intervocálica en el habla conversacional, si bien se han encontrado algunos casos de [j] en posición inicial absoluta en la pronunciación de las palabras en un contexto de frase, a menudo acompañada de un breve periodo cuasi-vocálico palatal -caracterizado por la aparición de formantes vocálicos de menor intensidad.

---

---

[w]	Aproximante labiovelar sonora
Características acústicas	Presencia de dos formantes muy próximos en la zona baja del espectro, de frecuencia inferior a los de la vocal [u].
Distribución	En lectura de frases, aparece en posición inicial absoluta seguida de [e], [i]. En el habla conversacional, aparece en posición intervocálica ante [e], [i], [o], y ante [a] como manifestación de un debilitamiento o fusión de [ɣ]+[u].

---

Las variantes [dʒ], [J], [j] y [j] se hallan en variación libre en posición inicial absoluta en una pronunciación cuidada. Su aparición no está condicionada por el timbre de la vocal siguiente ni por el acento, sino que responde básicamente a preferencias por parte del locutor: si bien la mayoría de informantes presenta un porcentaje mayor de aparición de [j], algunos muestran una clara tendencia hacia articulaciones oclusivas o aproximantes; la variante africada, por último, no aparece de forma generalizada sino sólo en ciertos locutores. La figura 1 muestra el número de casos de cada una de las variantes en función del informante.

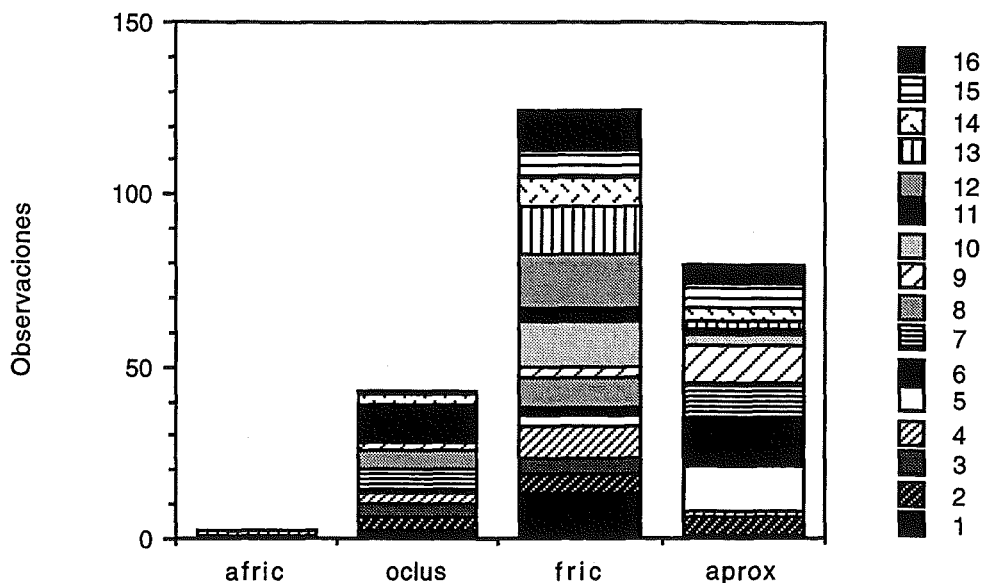


Figura 1. Número de observaciones de las variantes palatales africada, oclusiva, fricativa y aproximante en función del informante (1 a 16) en el corpus de palabras en frases marco.

En el habla conversacional, por otra parte, se hallan realizaciones fricativas [ʃ] y aproximantes [j] en contexto intervocálico, aunque con un predominio importante de las últimas en cualquier entorno vocálico y acentual; la variante fricativa se manifiesta como dependiente del informante, como se aprecia en la figura 2.

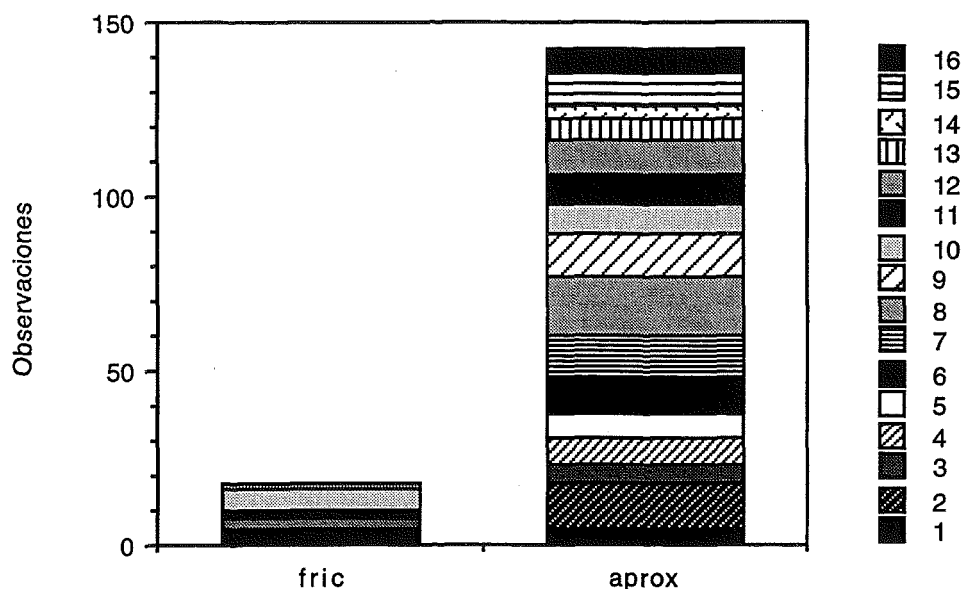


Figura 2. Número de observaciones de las variantes palatales fricativa y aproximante en función del informante (1 a 16) en el corpus de la tarea del mapa.

En la presente descripción pueden encontrarse ciertas discrepancias con las clásicas dedicadas al español (Navarro Tomás, 1918; Harris, 1969; Canellada-Kuhlmann, 1987).

Por un lado, la variante oclusiva sonora sólo aparece mencionada en Monroy Casas (1980), variante, que según sus observaciones, "utiliza el hablante de español académico"<sup>2</sup> si bien no aporta datos referidos a sus características fonéticas ni a su distribución en la lengua o a su frecuencia de aparición. Otro fenómeno que no ha sido destacado por ninguno de los autores consultados es la presencia de un periodo de nasalización previo a la realización de la consonante, en posición inicial absoluta, claramente distinto auditiva y acústicamente al periodo de sonoridad de las africadas.

<sup>2</sup> cf. nota 2, p. 83 y p. 98

Un problema adicional es la distinción entre las variantes fricativas [j] y las aproximantes [j̞]. En general, se ha considerado a la consonante española como fricativa palatal sonora (Navarro Tomás, 1918; Quilis, 1981; Canellada-Kuhlmann, 1987); sin embargo, otros autores (Martínez Celdrán, 1984b, 1991; Santagada-Gurlekian, 1989; Aguilar-Andreu, 1991) cuestionan dicha clasificación y ponen de manifiesto que las características acústicas del elemento no se corresponden con las de una fricativa- ruido de fricción- sino con las de una aproximante- estructura formántica definida-. Así, se sustituye el término de fricativa por el de aproximante para caracterizar a [j] junto a [β ð ɣ]. Sin embargo, en el presente análisis aparecen variantes fricativas [j̞] además de aproximantes [j], por lo que se considera necesario distinguir entre ambas. A partir de los resultados del análisis, podemos considerar a la realización aproximante como la característica del español en un contexto intervocálico si bien ciertos informantes prefieren la fricativa.

En cuanto a las variantes consonánticas velares, en el presente estudio se ha hallado una diferencia de manifestación de las mismas dependiendo de la situación de habla. En ambos conjuntos de datos, es posible aislar las siguientes realizaciones:

(1) una secuencia de oclusiva velar sonora caracterizada por la presencia de un segmento con las propiedades de una oclusiva sonora (sonoridad y explosión) seguido de una semiconsonante, segmento de corta duración, con los dos primeros formantes situados en la zona que corresponde a [u], identificable por la menor intensidad, la reducción de duración y la dirección de los formantes.

(2) una secuencia de aproximante velar sonora seguida de una semiconsonante. En este caso, la variación en la forma de onda y en la configuración formántica permite segmentar dos elementos; los formantes de la aproximante velar se muestran claramente influidos por las vocales adyacentes y, en cualquier caso, el F2 siempre aparece en una zona superior del espectro a la de [w],

(3) una aproximante labiovelar, caracterizada acústicamente por la presencia de dos formantes muy próximos en las frecuencias bajas, en una zona inferior a la de la vocal [u]. La principal diferencia entre una aproximante velar sonora y una aproximante labiovelar sonora se halla en la doble articulación de ésta última: el estrechamiento del canal simultáneo en la parte anterior y posterior causa un descenso de las frecuencias formánticas.

Sin embargo, la distribución varía en función de la situación de habla. En la lectura de frases, aparece una realización aproximante labiovelar en posición inicial absoluta ante [e], [i], y una oclusiva o aproximante velar sonora -en función de si aparece en posición inicial o en posición intervocálica- seguida de una semiconsonante -[ɣ̥ʷ]- ante [a] y [o]. Por el contrario, en el análisis de las muestras de habla conversacional, la aparición de [w] no puede relacionarse con el entorno vocálico, sino que es la variante característica ante [e], [i], [o], mientras que ante [a] se revela como resultado de un debilitamiento articulatorio: la agrupación [ɣ̥ʷ] ante [e i o] se manifiesta como [w] pero ante [a] se mantiene generalmente como [ɣ̥ʷa] excepto en ciertos casos de extrema relajación, en que aparece [wa]. En este sentido, hallamos un paralelismo con el proceso histórico según el cual los grupos QU, GU seguidos de A se mantuvieron en el paso del latín al romance frente a su fusión en cualquier otro contexto (Menéndez Pidal, 1904).

En este sentido, Navarro Tomás (1918) relaciona la variante [w] con las representaciones gráficas "hue", "hui", en posición intervocálica y en posición inicial absoluta; describe asimismo una variante [ɣ̥ʷ], manifestación relajada de [w] en una pronunciación familiar. La existencia de una fricativa velar labializada en español ha sido postulada por diferentes autores (Alonso, 1940; Malmberg, 1950, 1961; King, 1952; Bowen-Stockwell, 1955, 1956; Cárdenas, 1960; Hara, 1973) si bien desde presupuestos fonológicos, por lo que no se conocen sus especificaciones acústicas. En el presente trabajo, la relajación articulatoria se manifiesta en forma de un proceso [ɣ̥ʷ] > [w] y no a la inversa. No se ha podido establecer, por otra parte, la presencia de una fricativa velar labializada, sino bien una aproximante labiovelar, identificada acústicamente por la posición de sus formantes, bien la secuencia de una aproximante velar sonora seguida de una semiconsonante, reconocida por la variación de la forma de onda y la distribución frecuencial que se da en el paso de un segmento a otro. Así distinguimos entre las dos realizaciones, que en el habla conversacional se hallan en variación libre, contrariamente a Martínez Celdrán (1989) que opina que no existe diferencia fonética entre la u- de *hueso* y el grupo gu- de *guasa*, dado que se trata de un único fonema y un único sonido.

### 5.1.3. PROBLEMAS DE REPRESENTACIÓN

En el inicio del trabajo, se propuso una representación de las variantes vocálica, semiconsonántica, semivocálica y consonántica de acuerdo con las descripciones clásicas del español y tomando las posibilidades de simbolización que ofrece el Alfabeto Fonético Internacional (v. Tabla I).

DESCRIPCIÓN FONÉTICA	SÍMBOLO	DESCRIPCIÓN FONÉTICA	SÍMBOLO
vocal silábica cerrada anterior	i	vocal silábica cerrada posterior	u
vocal no silábica cerrada anterior (pre-vocálica y post-vocálica)	ɨ	vocal no silábica cerrada posterior (pre-vocálica y post-vocálica)	ɯ
aproximante palatal	j	aproximante labio velar	w
fricativa palatal sonora	ʃ	fricativa velar sonora labializada	ɣ <sup>w</sup>
africada palatal sonora	dʒ		

Tabla I. Descripción fonética del sonido y símbolo correspondientes según las directrices de la Asociación Fonética Internacional en su última versión del alfabeto (IPA, 1993).

Sin embargo, si queremos ser fieles a los hechos acústicos puestos de manifiesto a lo largo del trabajo, hay que reconsiderar la representación.

En la serie palatal, no se han hallado diferencias entre la variante asilábica en posición inicial del diptongo, la llamada semiconsonante, y la consonante palatal aproximante; sí aparecen diferencias, sin embargo, entre éstas y la variante asilábica en posición final de diptongo, o

semivocal. La semivocal presenta valores de F1 superiores y valores de F2 inferiores a los de la semiconsonante y a los de la vocal; se da por tanto una gradación de cierre y posterioridad vocal>semiconsonante>semivocal. Sin embargo, las diferencias sólo son significativas desde un punto de vista estadístico en F1 para semiconsonante vs. semivocal. Podemos interpretar que vocal y semiconsonante son más cerradas que la semivocal, pero la semiconsonante es menos anterior que la vocal. Por otro lado, si bien se hallan diferencias acústicas entre semiconsonante y vocal, éstas no aparecen entre semiconsonante y consonante aproximante.

En cuanto a la serie velar, encontramos distintas realizaciones acústicas correspondientes a la vocal, la semiconsonante, la semivocal y la consonante aproximante. La semivocal presenta valores de F1 y de F2 superiores a los de la semiconsonante y la vocal. Sin embargo, las diferencias no son estadísticamente importantes para F1; y en lo que respecta a F2, los indicios acústicos varían en función de la situación de habla. En la lectura de frases, hallamos diferencias de F2 entre semiconsonante vs. consonante, y entre semiconsonante vs. vocal, pero en el habla conversacional, no se dan diferencias semiconsonante vs. vocal, aunque es posible distinguir entre semiconsonante vs. consonante y consonante vs. vocal.

Esta situación podría hacernos pensar que en la lectura de listas, la supuesta consonante se ha pronunciado como vocal -debido a la influencia ejercida por la normativa- mientras que en el habla conversacional, el segmento adquiere su entidad consonántica. Por otro lado, en la lectura se observa la distinción vocal vs. semiconsonante, que desaparece en el habla conversacional: cuando el hablante deja de prestar atención sobre su lenguaje, la distinción semiconsonante vs. vocal se pierde, pero aflora la diferencia semiconsonante-vocal frente a consonante.

Hallamos así un factor de asimetría entre la serie palatal y la serie velar, en el sentido de que en la serie velar hay que diferenciar entre consonante y semiconsonante mientras que en la palatal estas dos entidades presentan comportamientos acústicos similares. Los procesos de asimilación aportan información en el mismo sentido: en la serie palatal, la semiconsonante y la consonante muestran un efecto parecido de palatalización sobre la vocal adyacente, pero en la serie velar la consonante velariza en un grado mayor a la vocal que la semiconsonante.





(c) atender únicamente a las características acústicas y no reparar en adoptar el símbolo procedente de las consonantes para la semiconsonante.

Dado que en la serie palatal, la consonante y la semiconsonante son similares acústicamente, y dado que en la serie velar, la semiconsonante se diferencia de la consonante por la falta de labialidad, obtendríamos la representación expuesta en la tabla IV.

semivocal palatal	[i̥+]	semivocal velar	[u̥+]
vocal anterior	[i]	vocal posterior	[u]
semiconsonante palatal	[j]	semiconsonante velar	[w]
consonante palatal	[j]	consonante velar	[w]

Tabla IV. Alternativa (c) de transcripción.

Adoptando la representación de semiconsonante y semivocal con el símbolo vocálico que alude a una variante menos extrema y más central que [i], [u] se aporta información sobre las características fonéticas de la realización y se evita la proliferación del uso de diacríticos: si además se desea expresar el carácter asilábico, puede adjuntarse el diacrítico de no silabicidad; así eliminaríamos la alternativa (a). Sin embargo, en español esta variante vocálica no aparece en entorno consonántico, contrariamente a otras lenguas -alemán, ruso- en que al valor fonético se suma un valor fonológico. Sería necesario por otra parte una comparación de las glides españolas con dichas vocales para poder establecer el uso del símbolo. Esta opción, además, no refleja las directrices de la Asociación Fonética Internacional según las cuales se evitará el uso de símbolos diferentes para aquellos sonidos que no causen distinciones fonémicas en la lengua (IPA, 1949; Wells, 1975). La decisión entre las alternativas planteadas comporta, por tanto, implicaciones sobre el sentido de una representación fonética y su relación con la interpretación fonológica del segmento. En último término, se trata de decidir si la simbolización fonética ha de basarse únicamente en información procedente del ámbito articulatorio, acústico y perceptivo, o bien ha de incorporar criterios fonológicos (cf. Beckman, 1990).

La alternativa (b), por ejemplo, presupone que la semiconsonante y la semivocal derivan de la vocal y que las diferencias entre semiconsonante y semivocal se explican por la dirección de la articulación; la alternativa (c), en cambio, implica que la semiconsonante es consonántica mientras que la semivocal es vocálica.<sup>3</sup>

En la discusión de la representación de las variantes consideradas, los argumentos que intervienen son básicamente tres:

- 1- la fidelidad a las propiedades acústicas,
- 2- la coherencia categorial entre los elementos del diptongo: semivocal y semiconsonante son elementos integrantes de un diptongo cuya principal diferencia es la posición pre- o post-vocálica,
- 3- la relación de alternancia entre la semiconsonante y la consonante en posición inicial de sílaba.

Teniendo en cuenta tales criterios, se obtienen las posibilidades de representación enumeradas a continuación. En primer lugar, para la serie palatal hallamos dos alternativas:

- 1.- Se mantiene la coherencia de representación en los elementos del diptongo, pero no responde fielmente a las características acústicas; tampoco se refleja la posibilidad de consonantización de la semiconsonante.

vocal anterior	[i]
semiconsonante palatal	[i <sub>+</sub> ]V
semivocal palatal	V[i <sub>+</sub> ]
consonante aproximante palatal	[j]

Tabla V. Alternativa de representación (1) en la serie palatal

<sup>3</sup>Ambas posturas han sido defendidas a lo largo de la historia de la fonología española (cf. cap. 3).

2- Se pierde la coherencia de representación en los elementos del diptongo, pero corresponde a los hechos acústicos observados; además, señala la alternancia semiconsonante-consonante en posición inicial absoluta.

vocal anterior	[i]
semiconsonante palatal	[j]V
semivocal palatal	V $\frac{[i+]}{i}$
consonante aproximante palatal	[j]

Tabla VI. Alternativa de representación (2) en la serie palatal.

En segundo lugar, para la serie velar, se dispone de tres opciones:

1- Se mantiene la coherencia de simbolización de los elementos del diptongo y se responde fielmente a las características acústicas, pero no queda reflejado el proceso de consonantización de la semiconsonante.

vocal posterior	[u]
semiconsonante velar	$\frac{[u+]}{u}$ V
semivocal velar	V $\frac{[u+]}{u}$
consonante aproximante labiovelar	[w]

Tabla VII. Alternativa de representación (2) en la serie velar.

2- Se pierde la coherencia de representación en los integrantes del diptongo además de la correspondencia con los hechos acústicos observados; sin embargo, alude al proceso de alternancia semiconsonante-consonante en posición inicial absoluta; mantiene además una simetría con la opción 2 de la serie palatal.

vocal posterior cerrada	[u]
semiconsonante velar	[w]V
semivocal velar	V[ $\frac{u}{\frac{1}{2}}$ ]
consonante aproximante labiovelar	[w]

Tabla VIII. Alternativa de representación (2) en la serie velar.

3- Se obtiene una correspondencia con las características acústicas de los sonidos, pero se pierde la relación entre los elementos del diptongo. En este sentido, para distinguir entre las tres realizaciones, hay que acudir al uso del diacrítico que indica "menos redondeado" junto al símbolo correspondiente a la consonante aproximante.

vocal posterior cerrada	[u]
semiconsonante velar	[w <sub>c</sub> ]V
semivocal velar	V[ $\frac{u}{\frac{1}{2}}$ ]
consonante aproximante labiovelar	[w]

Tabla IX. Alternativa de representación (3) en la serie velar.

Atendiendo a criterios estrictamente fonéticos, cada realización debería poder representarse con un símbolo propio, que aportaría información sobre sus características fonéticas y que sería independiente de la lengua. Sin embargo, el alfabeto propuesto por la Asociación Fonética Internacional no está basado en tal concepción, sino que tiene un enfoque fonológico (Ladefoged, 1990).

Aquí se opta por representar en el nivel fonético las variantes mediante los símbolos más adecuados de acuerdo con su caracterización acústica si bien pueden adivinarse las implicaciones fonológicas; así tenemos la propuesta de simbolización presentada en la tabla X.

DESCRIPCIÓN FONÉTICA	SÍMBOLO	DESCRIPCIÓN FONÉTICA	SÍMBOLO
vocal palatal	[i]	vocal velar	[u]
semiconsonante palatal	[j]V	semiconsonante velar	$[\frac{u}{\frac{1}{2}}+]V$
semivocal palatal	V $[\frac{i}{\frac{1}{2}}+]$	semivocal velar	$[\frac{u}{\frac{1}{2}}+]V$
consonante aproximante palatal	[j]	consonante aproximante labiovelar	[w]

Tabla X. Propuesta de transcripción.

En la serie palatal, semiconsonante y consonante se representan con el mismo símbolo dado que son acústicamente semejantes, mientras que en la serie velar, representamos a la semiconsonante como [u], basándonos en la vocal y no en la consonante por dos motivos básicos: por un lado, la diferencia frecuencial es menor entre vocal y semiconsonante, que entre semiconsonante y consonante, y por otro, en el habla conversacional, no se dan diferencias entre semiconsonante y vocal. Se adjuntan los diacríticos de avanzamiento y abertura para indicar que se trata de una variante más centralizada que  $[\frac{u}{\frac{1}{2}}+]$ .

En cuanto a la semivocal, las diferencias con respecto a la vocal se representan mediante la ayuda de diacríticos, dado que se considera que tales diferencias pueden explicarse por el movimiento articulatorio. Un diptongo se caracteriza por el movimiento rápido de una zona de frecuencias a otra, de ahí que el segundo elemento no llegue a los valores correspondientes a los de la vocal ideal: el diptongo [e̞i̞] se puede describir como la vocal [e] seguida del movimiento en dirección a [i], así como el diptongo [e̞ɥ̞] es la trayectoria desde la vocal [e] hacia las zonas de frecuencia de [u] sin alcanzar los valores ideales; de hecho, ninguno de los diptongos alcanza la posición de frecuencias indicada por los símbolos usados habitualmente en la transcripción.

#### 5.1.4. INVENTARIO FONÉTICO DEL ESPAÑOL

Las consideraciones expuestas en torno a las variantes vocálicas y consonánticas puestas de manifiesto en el experimento permiten plantear algunas cuestiones en torno al inventario fonético del español.

#### VOCALES

Según los resultados del presente análisis, se obtiene una representación del espacio vocálico como se ofrece en la figura 1, donde se incluyen además las variantes abiertas y cerradas de las vocales medias [e] y [o], tal como describe Navarro Tomás (1916).

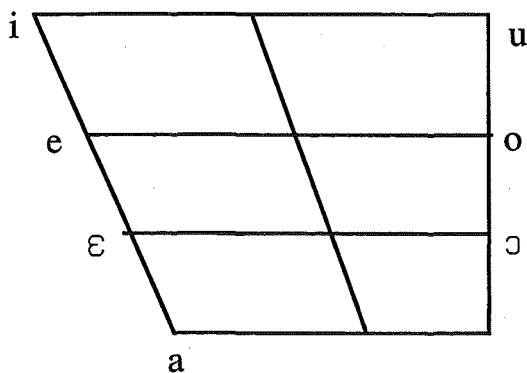


Figura 1. Inventario vocálico del español.

La figura 2, por su parte, presenta el conjunto de diptongos del español. Las flechas indican la dirección del movimiento frecuencial, y el final de las mismas la zona de frecuencias alcanzada.



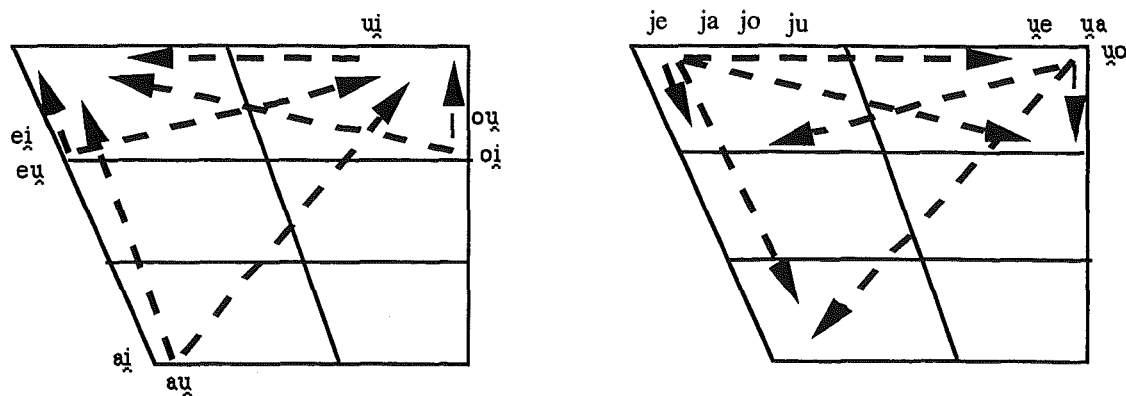


Figura 2. Conjunto de diptongos del español.

## CONSONANTES

Si tomamos las descripciones clásicas de Navarro Tomás (1918) y Canellada-Kuhlmann (1987), surgen diferentes problemas de caracterización de sonidos a la vez que si nos interesa representar los sonidos del español atendiendo a los criterios de la Asociación Fonética Internacional, se presentan problemas de simbolización.

En cuanto al inventario, según los resultados del experimento habría que considerar la aparición de una oclusiva palatal sonora y una fricativa palatal sonora, además de incorporar la aproximante palatal sonora.

Por lo que se refiere a la representación, si nos basamos en el Alfabeto Fonético Internacional en su última versión de 1993, hay que reconsiderar el uso de los símbolos para el español. En de la Mota - Ríos (1993) se discute el inventario de símbolos expuestos en la figura 3.

	Bilabial	Labio dental	Inter dental	Dental	Alveolar	Post-alveolar	Retroflejo	Palatal	Velar	Uvular	Faringal	Glotal
Oclusivo	p	b	t <sup>+</sup>	t <sup>h</sup> d <sup>h</sup>					k g			
Nasal	m	ɱ	n <sup>h</sup>	n <sup>h</sup>	n			ɲ		ŋ		
Vibrante					r							
Flap												
Fricativo		f	θ θ <sup>h</sup>	s s <sup>h</sup>	s s <sup>h</sup>			ç x				
Lateral (fricativo)												
Aproximante	β β <sup>h</sup>				ɹ							
Lateral (Aproximante)			l <sup>h</sup>	l <sup>h</sup>	l			ʎ				
Eyectivo												
Implosivo												
Africado							ʧ ʧ <sup>h</sup>					

Figura 3. Adaptación del AFI (versión 1989) al español según de la Mota-Ríos (1993).