

## 3 Caracterización acústica de las principales clases de sonidos del español

### 3.1 Introducción

El valor comunicativo de los gestos articulatorios que hemos descrito en el capítulo previo descansa primariamente en las distintas maneras en que la posición de los articuladores modifica el paso del aire, produciendo ondas sonoras diferentes.<sup>1</sup> Estas ondas sonoras, rápidos cambios en la presión del aire propagados como olas en un estanque, impactan los órganos auditivos del oyente, quien los interpretará como portadores de significado si conoce la lengua.

En este libro, los sonidos del español se describen sobre todo desde una perspectiva articulatoria. Sin embargo, en este capítulo haremos referencia también a las características acústicas de los principales tipos de sonidos del habla. Mientras que en el pasado el estudio acústico del habla requería aparatos complejos y caros, hoy en día existen programas de *software* de bajo precio o incluso gratuitos que los lectores de este libro pueden utilizar para analizar su propia pronunciación y otras grabaciones. Estos programas de análisis del habla proporcionan una representación visual de las ondas sonoras y permiten varias operaciones de análisis y filtrado. Todas las figuras de este capítulo y la mayoría del libro han sido producidas con el programa Praat (desarrollado por Paul Boersma y David Weenink, [www.praat.org](http://www.praat.org); véase Boersma 2001).

En las siguientes secciones señalaremos las principales características visuales que nos permiten identificar en un espectrograma las distintas clases de segmentos. En otros capítulos se dará algo más de información sobre las características acústicas de sonidos específicos.<sup>2</sup>

<sup>1</sup> Primariamente pero no completamente. La información visual que nos da el movimiento de los labios es también muy importante en la percepción, como demuestra el llamado “efecto McGurk” (McGurk y MacDonald 1976).

<sup>2</sup> Una exposición detallada de fonética acústica está fuera de los límites marcados para este libro. Algunas referencias excelentes son Laver (1994), Ladefoged (1992, 1996), Johnson (2003) y Hayward (2000). Para el análisis espectrográfico de los sonidos del español véase Quilis (1981) y Martínez

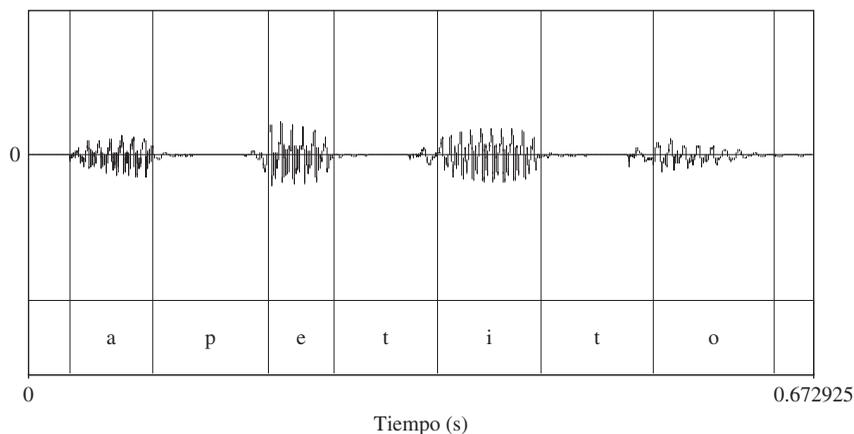


Fig. 3.1 Onda sonora (oscilograma) de una producción de la palabra *apetito*

### 3.2 Vocales y oclusivas sordas

Dos tipos de sonidos del habla máximamente diferentes entre sí son las vocales y las oclusivas sordas. En la figura 3.1 podemos observar la onda sonora (el oscilograma) de una producción de la palabra *apetito* /apetító/. En esta representación, la dimensión horizontal corresponde al tiempo (en segundos), mientras que la amplitud de la onda (su desplazamiento vertical con respecto a la línea cero, que representa la presión normal del aire) muestra la cantidad de energía o intensidad a cada momento (aumento o disminución de presión, por encima o debajo de la línea cero, respectivamente). La amplitud general de la onda dependerá de lo fuerte que hablemos (o lo cerca que tengamos el micrófono), pero los diferentes sonidos también difieren en su amplitud relativa. En general, las vocales producen ondas con mayor amplitud que los segmentos colindantes.

Observemos primero en la figura 3.1 las secciones correspondientes a las tres oclusivas sordas. En la articulación de estas consonantes los órganos articulatorios hacen contacto completo bloqueando totalmente el paso del aire por un instante. En la figura, este momento de silencio se refleja en la ausencia de energía durante la fase de oclusión de las tres oclusivas.

En la producción de las vocales, como en las de otros sonidos sonoros, la corriente de aire que sale de los pulmones causa la vibración de las cuerdas vocales, es decir, que se abran y cierren rápidamente (véase la sección 2.2.3). En la articulación de las vocales la corriente de aire pasa sin obstáculo por la

Celdrán (1998). Para el inglés americano, lengua a la que nos referiremos en varias ocasiones, véase Olive et al. (1993).

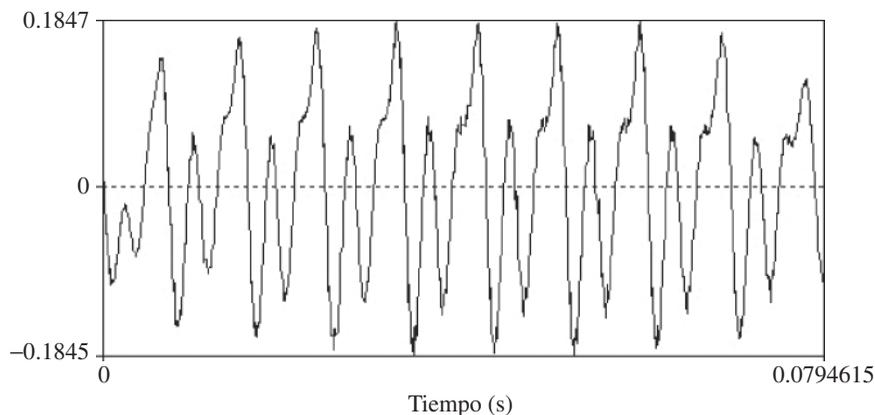


Fig. 3.2 Onda cuasi-periódica de /i/ en *apetito*

cavidad oral y es modulada de diversas maneras según el tamaño de las cavidades anterior y posterior creadas por la posición de la lengua.

La vibración compleja de las cuerdas vocales puede compararse con la producida cuando tocamos la cuerda de una guitarra. En la figura 3.1 las secciones correspondientes a las vocales son observables como ondas cuasi-periódicas; esto es, como ondas con un patrón que se repite, en las que cada repetición corresponde a una vibración de las cuerdas. Esto se puede ver mejor en la figura 3.2, que muestra solo el oscilograma de la tercera vocal en la figura 3.1, /i/.

El número de repeticiones o ciclos por unidad de tiempo de una onda periódica se conoce como su frecuencia. La frecuencia se da en ciclos por segundo, *hertz* o *hercios* (abreviado Hz). En la figura 3.3, tenemos la onda de otra vocal, /a/. La duración de este segmento aparece en la parte inferior de la figura como 0.100175 segundos, o apenas algo más de una décima de segundo. Si contamos los picos de amplitud en la figura podemos ver que la onda se repite exactamente diez veces en este intervalo de tiempo. En un segundo habría pues 100 ciclos. Decimos entonces que la frecuencia fundamental de esta onda es de unos 100 ciclos por segundo o 100 Hz. Las diferencias de frecuencia fundamental se trasladan en diferencia de tono percibido. Cuanto más alta es la frecuencia fundamental de una onda, más alto es el tono que oímos. La frecuencia fundamental de una onda periódica depende de la velocidad de vibración o apertura y cierre de las cuerdas vocales.

En la figura 3.2 y la figura 3.3 podemos ver que sobrepuestos sobre el ciclo de mayor amplitud (la frecuencia fundamental) hay otros componentes periódicos de frecuencias más altas. Estos componentes se conocen como armónicos y son siempre múltiplos íntegros de la frecuencia fundamental. La amplitud de los armónicos por encima de la frecuencia fundamental revela de qué vocal se trata.

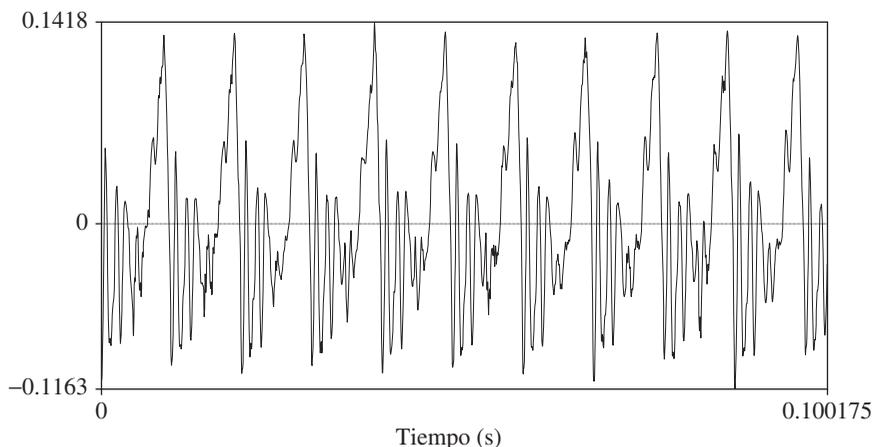


Fig. 3.3 Oscilograma de la vocal /a/ producida con una frecuencia fundamental de unos 100 Hz. En la figura hay 100 ciclos en una décima de segundo

Nótese que las vocales /i/ (figura 3.2) y /a/ (figura 3.3) tienen diferentes patrones de armónicos, que aparecen como picos y valles en cada ciclo.

Además de la información que obtenemos mediante la inspección visual de ondas sonoras, podemos obtener información adicional sobre los sonidos del habla mediante el análisis de los espectrogramas. Un espectrograma es una representación de la señal acústica que nos da más detalle al permitirnos observar la distribución de la energía a diferentes frecuencias. Esto nos permite, por ejemplo, distinguir más claramente entre una vocal y otra. La figura 3.4 es un espectrograma producido a partir de la onda sonora en la figura 3.1. En un espectrograma, como en una onda sonora, el eje horizontal indica el tiempo. El eje vertical, por otra parte, indica diferentes frecuencias (en la figura 3.4, por ejemplo, desde 0 Hz a 5000 Hz) y un color más oscuro indica mayor intensidad en las frecuencias correspondientes.

En la figura 3.4 podemos ver que la energía de las vocales se concentra en ciertas bandas de frecuencias (en las zonas más oscuras). Las bandas horizontales más oscuras se conocen como **formantes**. Cada formante es un haz de armónicos. Como veremos en el capítulo 6, cada vocal tiene una distribución de formantes característica. Por ejemplo, el primer formante es más alto para /a/ que para cualquier otra vocal.

Para las oclusivas de /apetito/ en la figura 3.4, podemos observar que el espectrograma está completamente en blanco durante la oclusión. El suelte o explosión de la oclusión consonántica queda reflejado en una línea vertical seguida de un corto intervalo de energía aperiódica (sin un patrón reiterado definido) o “ruido”, antes de comenzar los formantes de la vocal siguiente.

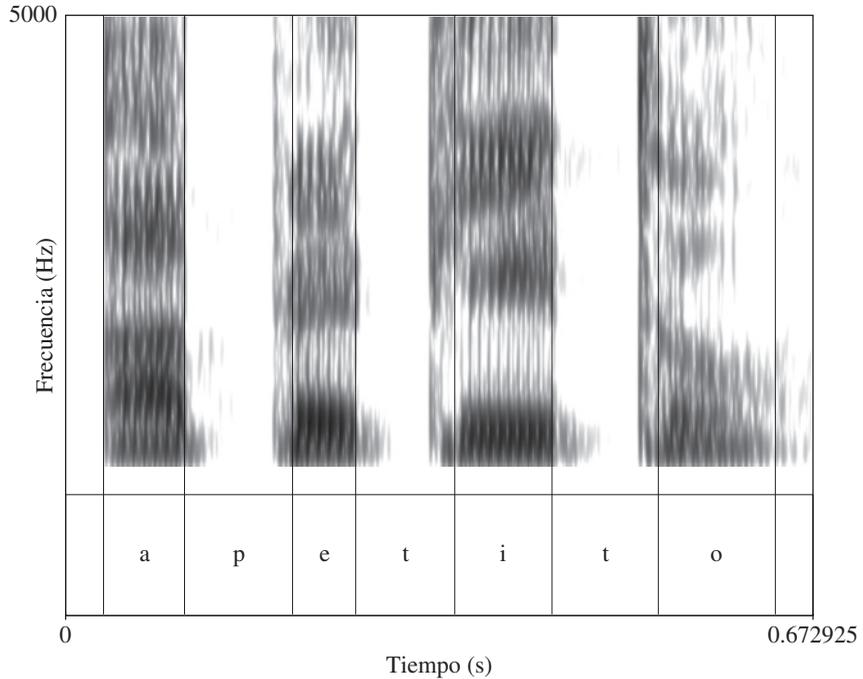


Fig. 3.4

Espectrograma de *apetito*. Las vocales se caracterizan por la presencia de formantes (barras horizontales más oscuras). Cada oclusiva sorda tiene un período de silencio (en blanco en el espectrograma) seguido de energía no periódica durante la explosión

El centro de energía en la barra vertical, así como la forma de las transiciones de los formantes de las vocales adyacentes, nos dan información sobre el punto de articulación de la consonante.

### 3.3 Fricativas y africadas

La “firma visual” de las fricativas es muy diferente tanto de la de las oclusivas como de la de las vocales. En las fricativas sordas, la energía acústica no tiene su fuente en la vibración de las cuerdas vocales, sino en el estrechamiento del paso del aire creado por los articuladores en la cavidad oral, que produce turbulencia. Las fricativas tienen una onda sonora no repetitiva o aperiódica, la posición de cuyo centro de energía (de la zona más oscura en el espectrograma) en frecuencias más altas o más bajas depende de su punto de articulación. En la figura 3.5 tenemos la onda sonora de una producción de la palabra *ositos* /osítos/. Obsérvese la gran diferencia que hay entre las secciones de la onda sonora de esta palabra que corresponden a los dos casos de la fricativa /s/ y las porciones que corresponden a las vocales y a la oclusiva /t/.

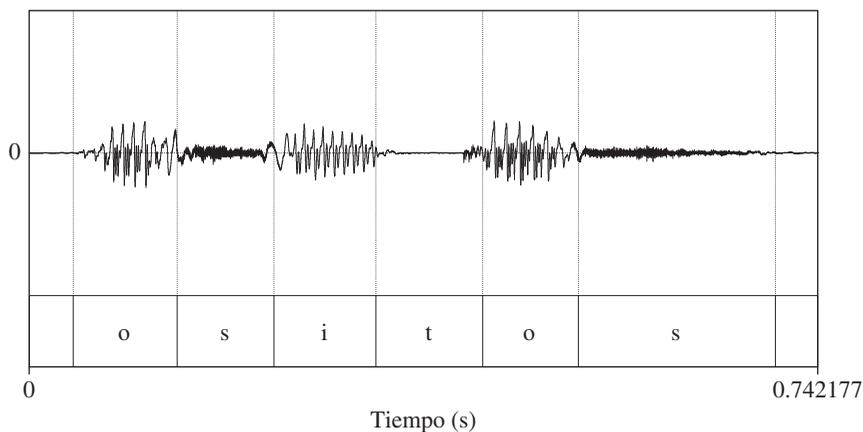


Fig. 3.5 Oscilograma de una producción de la palabra *ositos*

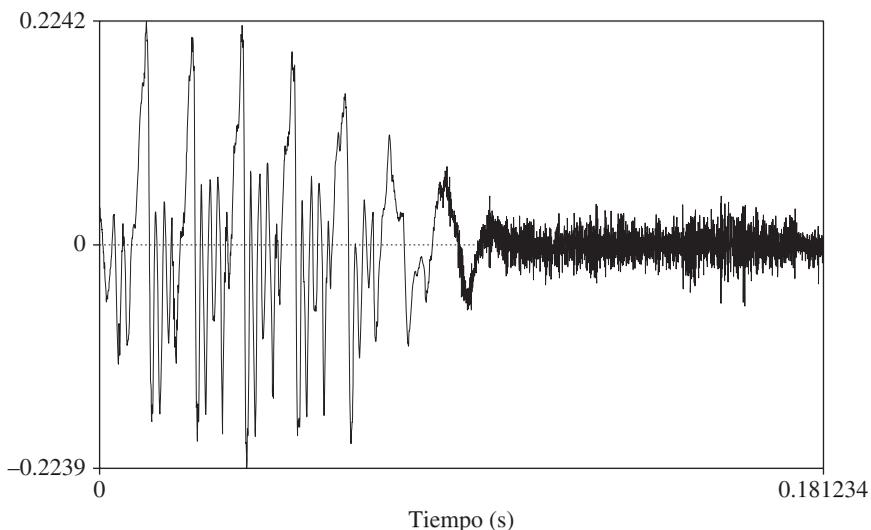


Fig. 3.6 Oscilograma de /-os/ de la fig. 3.5

En la figura 3.6 hemos ampliado parte de los dos últimos segmentos, /-os/, de modo que pueda verse de manera aún más clara la diferencia que hay entre la onda periódica de la vocal /o/ y la onda aperiódica (sin patrón definido) de la fricativa /s/.

La figura 3.7 es un espectrograma de /ositos/ obtenido de la misma señal que la figura 3.5. Nótese que para /s/ la energía se distribuye de manera no organizada, sin formantes, en la parte superior del espectrograma. El oscilograma aparece también en la parte superior de la figura, por encima del espectrograma.

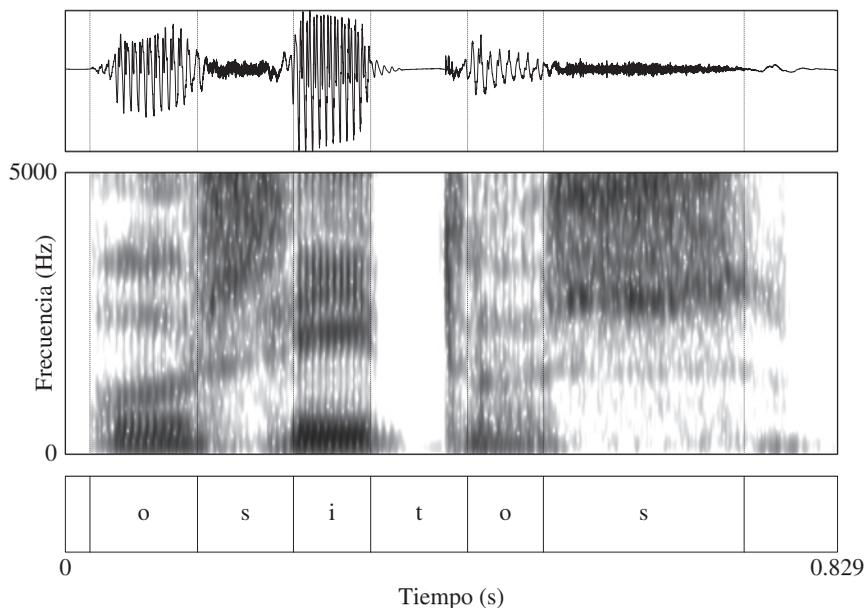


Fig. 3.7 Oscilograma y espectrograma de /osítos/

Las africadas combinan oclusión con suelto fricativo. Las dos fases (el momento de silencio, seguido de fricción similar a la de la fricativa /s/ en la figura 3.7) pueden verse en el espectrograma de *hacha /áʃa/* en la figura 3.8.

### 3.4 Oclusivas sonoras y alófonos aproximantes de /b d g/

Los fonemas oclusivos sonoros de /b d g/ en español, como sabemos, se realizan sin oclusión completa en muchas posiciones, incluyendo la intervocálica. Su grado de constricción puede variar considerablemente. Cuanto más abierta sea su articulación, más se parecerá su espectrograma al de las vocales y más difícil será encontrar en el espectrograma las fronteras entre consonante y vocales colindantes. En la figura 3.9 tenemos un espectrograma de una producción de la frase *sabe todo /sábe tódo/* [sáβe tóðo].

Las características espectrográficas de las oclusivas sonoras en posición intervocálica pueden ser ilustradas con un ejemplo del inglés, dado que normalmente no ocurren en español. En la figura 3.10 se comparan una producción del español *ave /ábe/*, que contiene una aproximante sonora [áβe], con una producción de la palabra inglesa *abbey* ‘abadía’ pronunciada por un anglohablante. Nótese que en la /b/ del ejemplo inglés la única energía observable aparece en la parte inferior del espectrograma. Esto es lo que se conoce como **banda o barra de sonoridad** e indica que el segmento es sonoro. Todos los

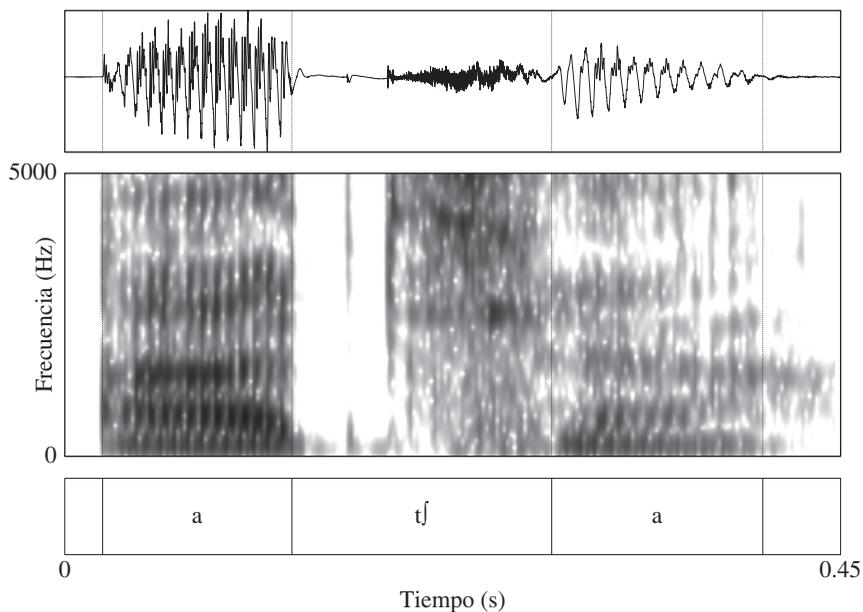


Fig. 3.8 Oscilograma y espectrograma de una producción de *hacha* /âƒa/. Obsérvense las dos fases durante la producción de la africada

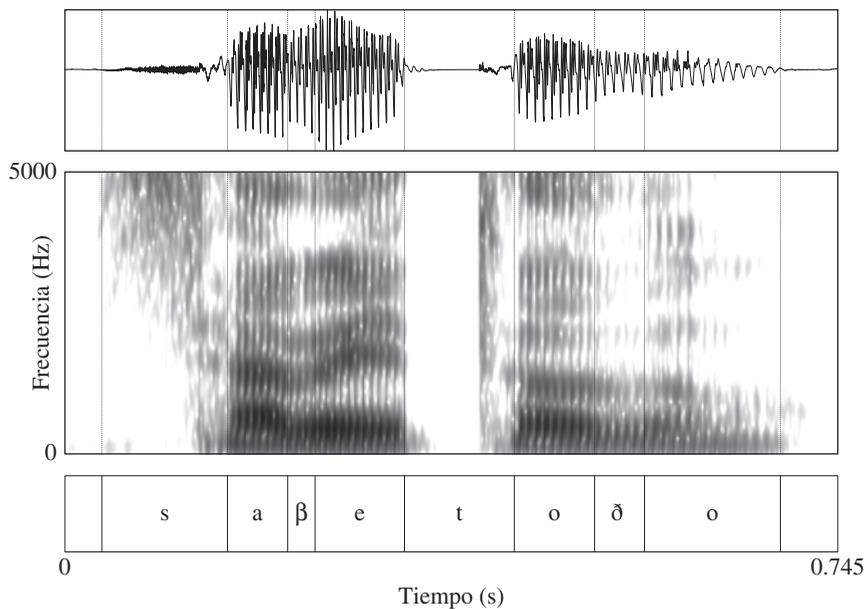


Fig. 3.9 Oscilograma y espectrograma de una producción de *sabe todo*

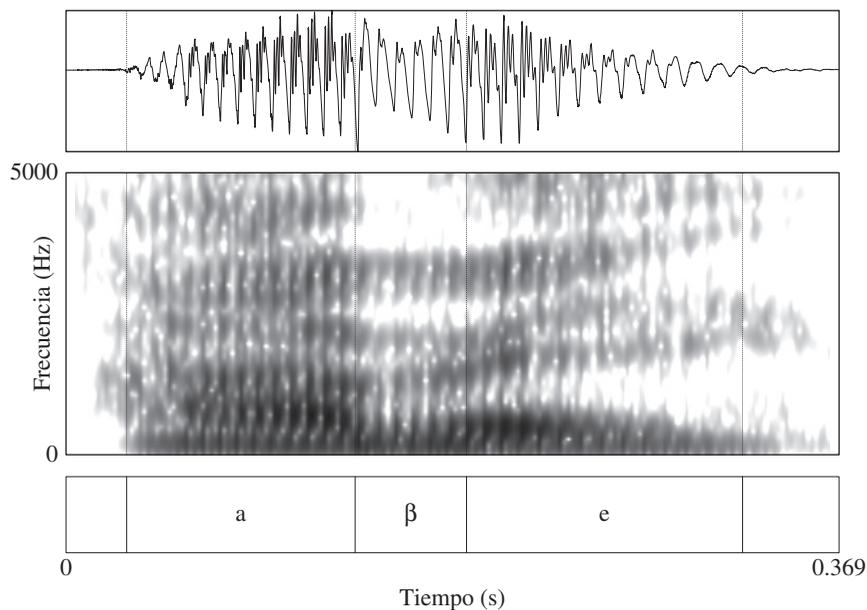


Fig. 3.10a Oscilograma y espectrograma de una producción de *ave* /ábe/ [áβe]

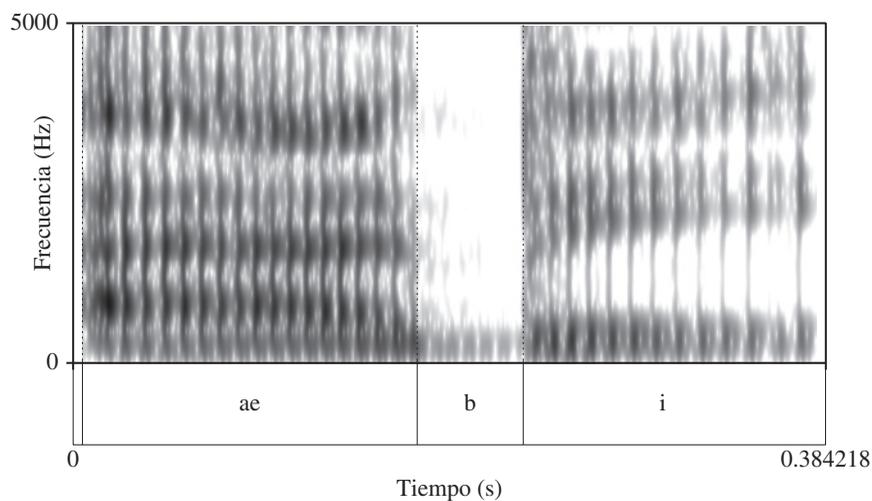


Fig. 3.10b Espectrograma de una producción de la palabra inglesa *abbey* /'æbi/ 'abadía'

segmentos sonoros tienen esta banda de energía a bajas frecuencias. En la /b/ del inglés, que ha sido realizada como oclusiva sonora, esta es la única energía visible en el espectrograma. En el ejemplo del español, por otra parte, el fonema /b/ ha sido realizado como la aproximante [β] y esto se refleja en la presencia de formantes como los de las vocales colindantes, si bien algo menos intensos.

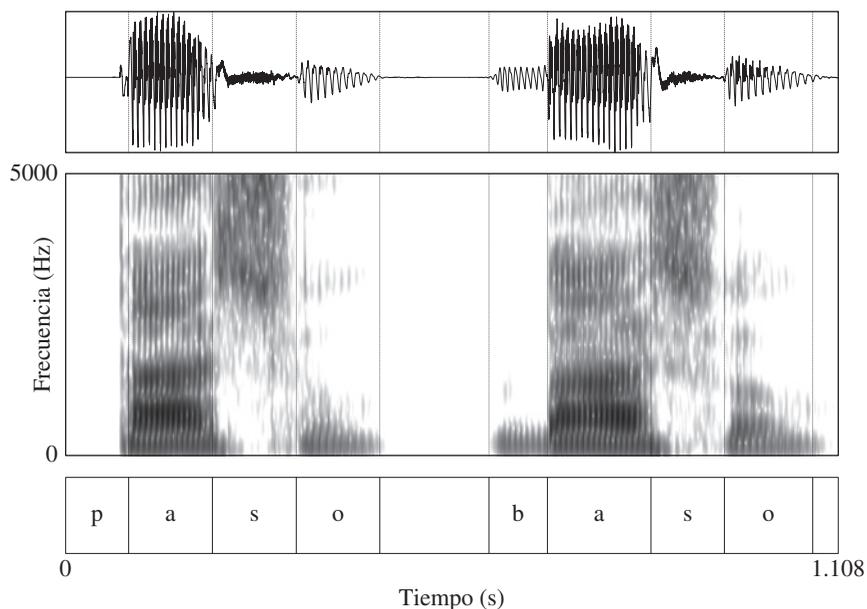


Fig. 3.11 Oscilograma y espectrograma de una producción de las palabras *paso* /páso/ y *vaso* /báso/

En la figura 3.11 se comparan *paso* /páso/ y *vaso* /báso/. En posición inicial absoluta, a principio de enunciado, /b/ es una oclusiva. Nótese que /b/ tiene banda de sonoridad, que indica vibración durante la oclusión (como la /b/ de *abbey* en la figura 3.10b), mientras que esta banda de sonoridad no aparece en /p/. Por otra parte, la apertura de /p/ se marca por una breve aspiración (mucho más breve que la de la consonante /p/ del inglés en la misma posición).

### 3.5 Consonantes resonantes

El breve contacto del ápice en la vibrante simple resulta en un espectrograma similar al de una oclusiva sonora, pero muy breve. La secuencia de breves contactos en la vibrante múltiple produce una serie de espacios en blanco (excepto por la banda de sonoridad) en el espectrograma. Las dos vibrantes, simple y múltiple, se ilustran en la figura 3.12 con las palabras *para* /pára/ y *parra* /párra/. En este caso, la vibrante múltiple de *parra* ha sido producida con tres contactos.

Las nasales y laterales también tienen formantes, aunque menos intensos que los de las vocales adyacentes. En la figura 3.13 ofrecemos el espectrograma de una producción de *la lana* /lalána/.

En otros capítulos daremos algún detalle más sobre las características acústicas de cada sonido o clase de sonidos, según los vayamos estudiando.

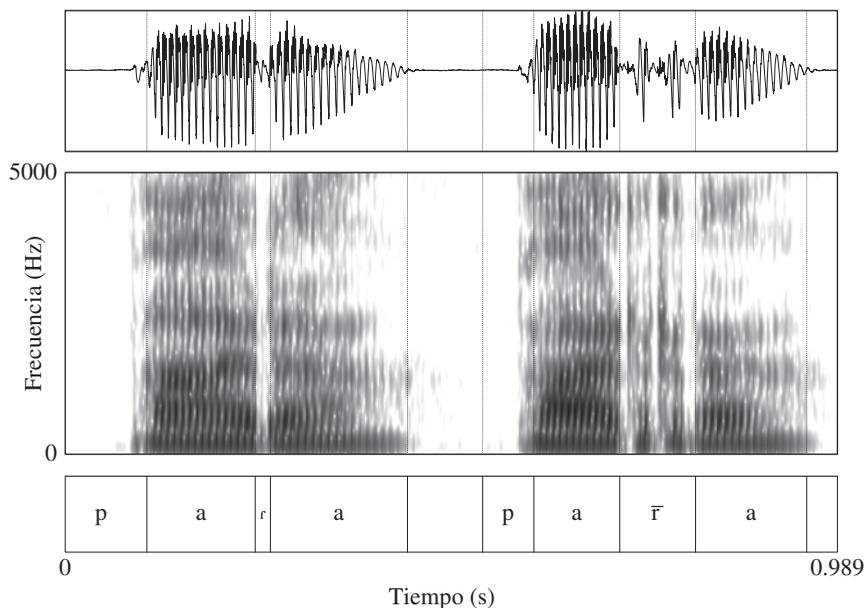


Fig. 3.12 Oscilograma y espectrograma de una producción de las palabras *para* /pára/ y *parra* /pâra/. Obsérvese la breve oclusión intervocálica en *para* y las tres oclusiones en esta producción de *parra*

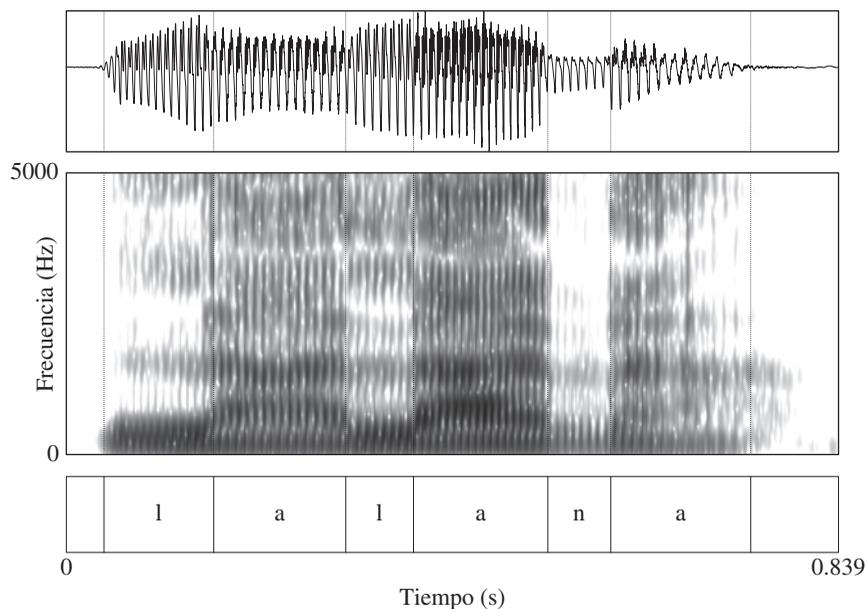


Fig. 3.13 Oscilograma y espectrograma de una producción de *la lana*