

La fonética acústica 2: las consonantes

Dr. Christian DiCano
cdicano@buffalo.edu

Universidad de Búfalo

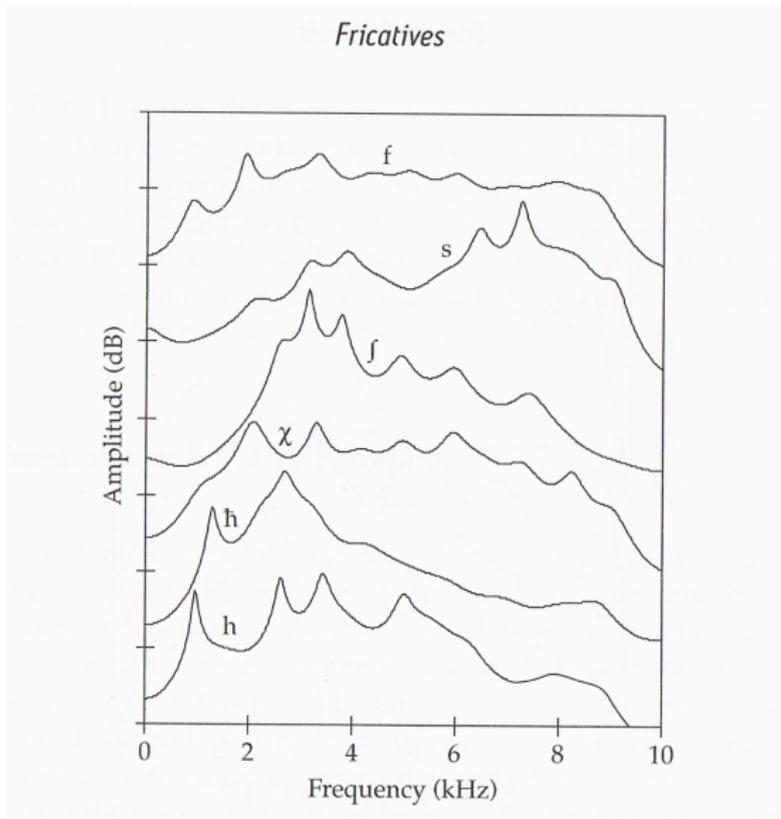
21/6/18

Cómo se genera el sonido en las fricativas?

La fuente del sonido es la turbulencia creada en el punto de constricción.

Hay dos lugares con constricción en muchas fricativas: la fuente y los obstáculos más adelante en la boca (los labios, los dientes).

El ruido es aperiódico pero está filtrado por las resonancias en la boca y en los obstáculos. Entonces unos rangos de frecuencia tienen más amplitud con fricativas diferentes. Hay fricativas con energía más alta y más baja.



(Johnson, 2012:163)

La ubicación de las cimas del espectro de una fricativa se predice por la longitud del espacio enfrente de la constricción.

Si el espacio es más largo, la cima será más bajo (como F2). Si el espacio es menos largo, la cima será más alta.

Fricativas anteriores tienen frecuencias más altas de fricativas posteriores.

Ejemplo de Praat - fricativas de inglés.

Fricativas labiales y interdientales

- Algunas fricativas no tienen un espacio enfrente de la constricción (β , ϕ , θ , δ , f , v). Cómo parecen en la señal?
- El espectro es más plano con menos cimas (no hay resonancias). Resulta que pocos idiomas distingue entre $/\beta/$ y $/v/$, o entre $/\phi/$ y $/\theta/$.
- Fricativas dentales/interdentales parecen mucho a las fricativas labiodentales por sus propiedades espectrales. Se los distingue más por sus propiedades transicionales.
- Es decir, el oyente usa más la transición de los formantes previos o siguientes para distinguir estas fricativas. (Ejemplo en Praat)

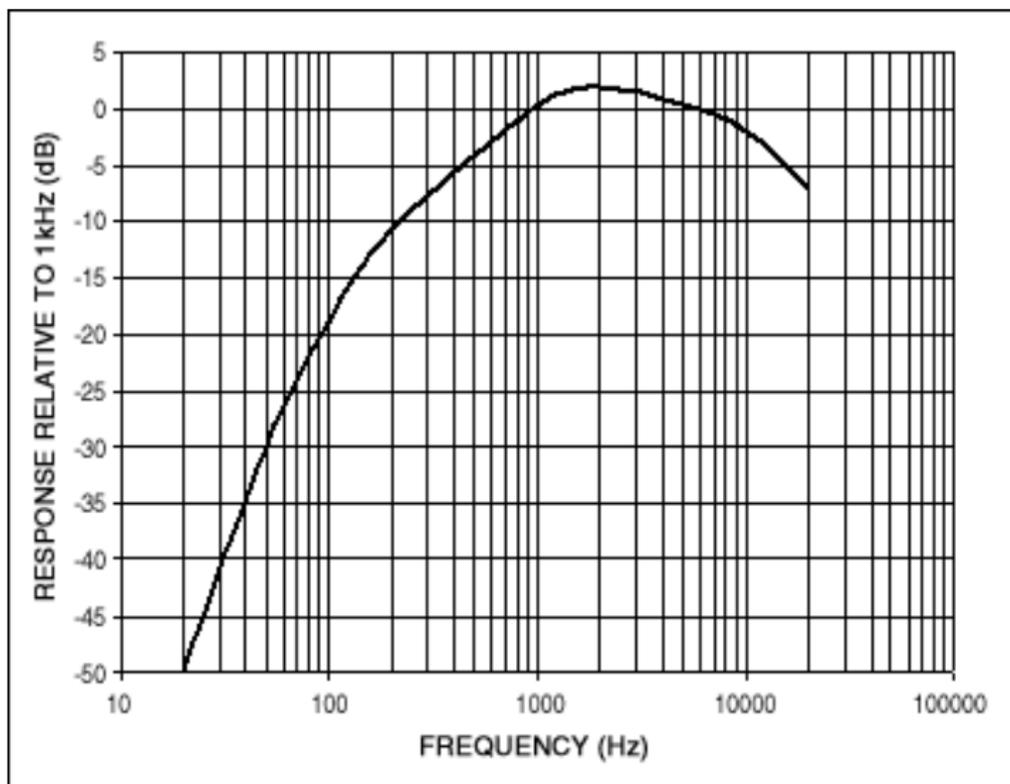
Sibilantes

- Se caracteriza los sibilantes por un canal central que dirige el aire directamente y rápidamente a los dientes.
- Una distinción entre [s] y [ʃ] es la longitud de la cavidad; [s] tiene componentes de frecuencia más altos que [ʃ].
- Sin embargo muchas veces se exagera la distinción entre estos sonidos por el redondeamiento de los labios con [ʃ]. Eso baja la concentración de frecuencias aún más y produce una cavidad sublingual que extiende la cavidad enfrente de la constricción.
- La concentración de energía espectral se queda entre 3500-4000 Hz y 6000-10000 Hz con [s] y entre 2500-3500 Hz con [ʃ]; (véase Praat).

Si la diferencia en la frecuencia es muy larga entre estos sibilantes, por qué se debe exagerar la diferencia?

- Lo que vemos en la señal acústica no es lo que oyemos.
- Discriminamos entre frecuencias muy altas peor que entre frecuencias más bajas. Entonces las frecuencias más altas suenan más parecidas a nuestros oídos.
- Si examinamos los espectros en esta manera 4000 vs. 6000 Hz. no son muy diferentes.

Eso también explica por qué no usamos un límite de frecuencia más de 10000 Hz para sonidos de habla; no podemos distinguir entre estas frecuencias altas muy bien.



Fricativas posteriores

- Las fricativas más posteriores tienen más cimas por el tamaño de la cavidad enfrente de la constricción.
- La ubicación de las cimas parece la ubicación de los formantes vocálicos para vocales posteriores.
- Hay más coarticulación entre fricativas posteriores y las vocales adyacentes que observamos con fricativas anteriores.

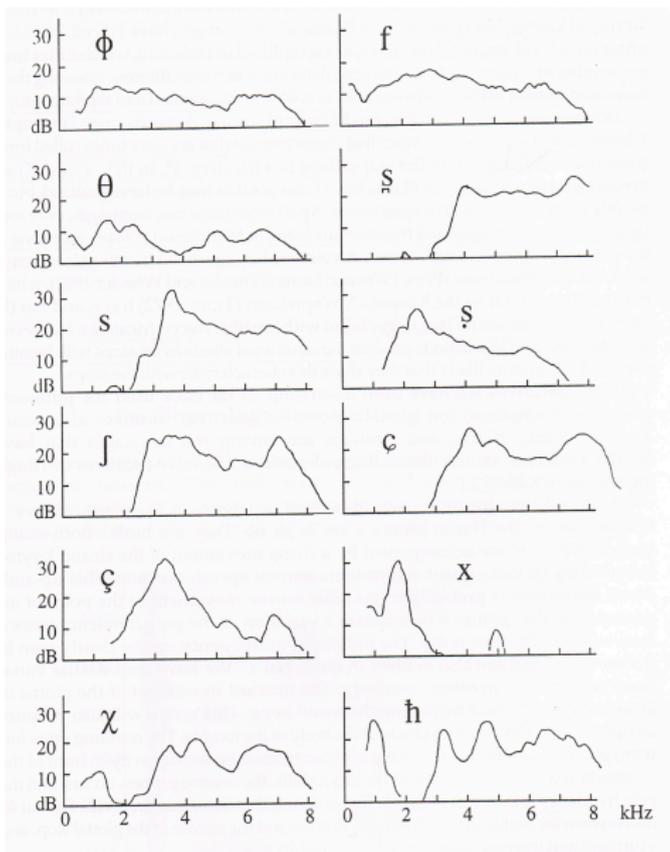
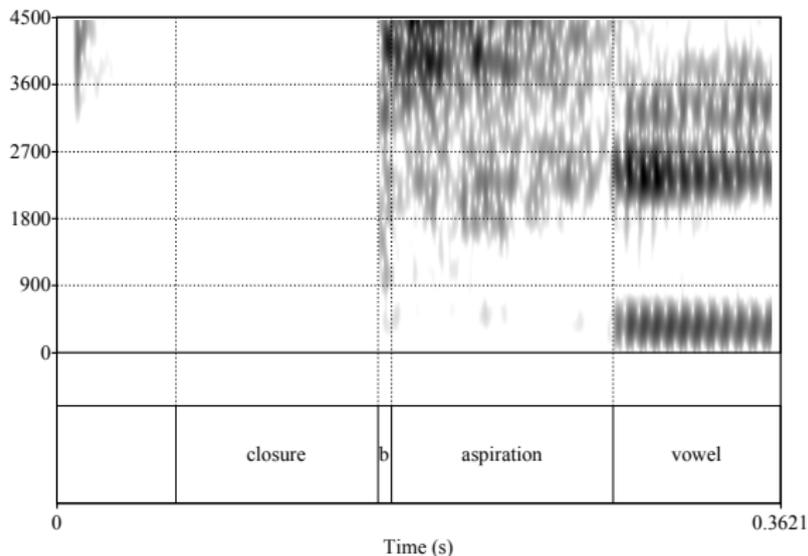


Figure 5.3. Spectra of 12 fricatives as produced by Jassem (1968).

Oclusivas y africadas

- En la señal acústica nos interesan tres componentes de la producción de las oclusivas: el inicio del gesto de cerrar, el cierre, y la apertura.
- Las oclusivas sordas y aspiradas se produce con una desplegación del glotis donde las cuerdas vocales están la más abiertas cuando empieza la apertura.
- La duración de la apertura de oclusivas aspiradas contiene dos eventos acústicos distintos: una ruptura producida en la ubicación de la constricción y el ruido de aspiración producida en el glotis.

La palabra [t^him] 'equipo.'



La ruptura tiene energía dispersa a través del espectro pero con una concentración de energía cerca de las resonancias de la porción anterior de la cavidad oral.

VOT

El VOT (voice onset time - tiempo de apertura de sonorización) es la demora entre la ruptura de una oclusiva y el comienzo de la sonorización del segmento siguiente.

Es muy corto con oclusivas sordas no aspiradas, más largo con oclusivas aspiradas, y negativo con oclusivas sonoras.

- Típicamente la duración de la ruptura varía con el punto de articulación de la oclusiva. Oclusivas anteriores tienen rupturas cortas y oclusivas posteriores tienen rupturas largas. Se surge del tiempo requerido para mover articuladores de tamaños diferentes (Stevens, 2000).
- Sin embargo, todas las rupturas son acústicamente cortas: entre 2-6 ms para oclusivas bilabiales y alveolares; entre 10-25 ms para oclusivas velares.
- El amplitud de la ruptura se determina por la diferencia en presión a través del punto de constricción. Es muy fuerte con clicks y ejectives y más débil (con menos intensidad) con oclusivas pulmonicas. Es lo más débil con implosivas.

VOT y el punto de articulación

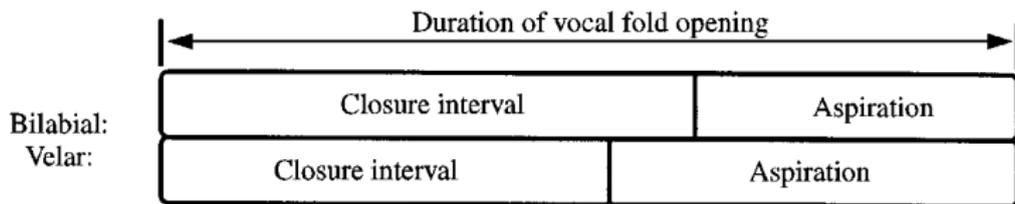
TABLE IV. Mean VOT (ms) of the stops in 18 languages studied in the UCLA endangered languages project (as of May 1999)

Language	Bilabial	Dental	Alveolar	Retroflex	Velar	Uvular
Aleut (Eastern)			59		75	78
Aleut (Western)			76		95	92
Apache	13		15		31	
Apache (aspirated)			58		80	
Banawá		22			44	
Bowiri	17		18		39	
Chickasaw	13		22		36	
Dahalo	20	15	42		27	
Defaka	18		20		30	
Gaelic	13	22			28	
Gaelic (aspirated)	64	65			73	
Hupa	11	16			44	27
Hupa (aspirated)		82			84	
Jalapa Mazatec			11		23	
Jalapa Mazatec (aspirated)			63		80	
Khonoma Angami	10	9			20	
Khonoma Angami (aspirated)	83	55			91	
Montana Salish	22		24		48	55
Navajo	12		6		45	
Navajo (aspirated)			130		154	
Tlingit			18		28	30
Tlingit (aspirated)		120			128	128
Tsou	11		17		28	
Wari'	19	26			50-58	
Yapese	20	22			56	

(Cho and Ladefoged, 1999)

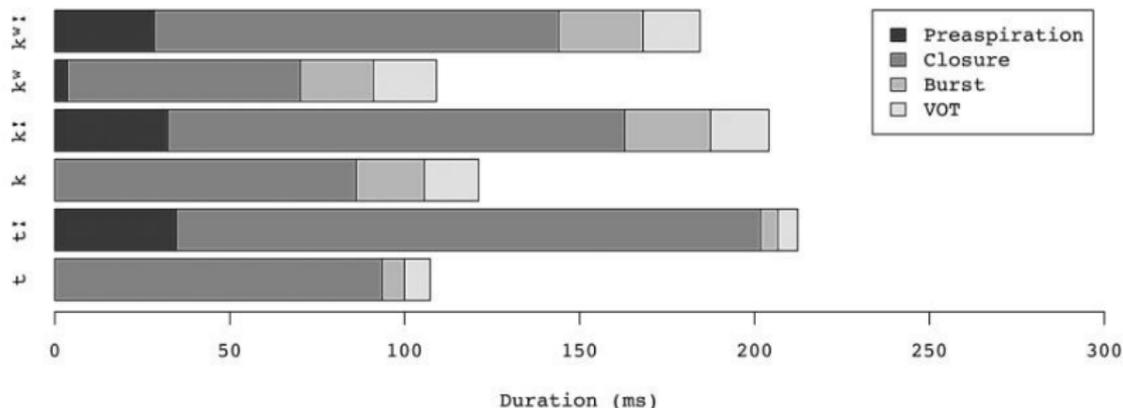
Relaciones de intercambio y VOT

Las rupturas y la duración de aspiración son más largas con oclusivas posteriores que con oclusivas anteriores. Pero a la misma vez, la duración del cierre de oclusivas posteriores es más corta que lo que observamos con oclusivas anteriores.



Sugiere que haya una relación de intercambio de duración (Maddieson, 1997). La duración entera de la oclusiva se determina por el cronometraje del glotis pero la duración de los componentes (el cierre, la ruptura) pueden variar.

Evidencia de Triqui de Itunyoso (DiCanio, 2012)

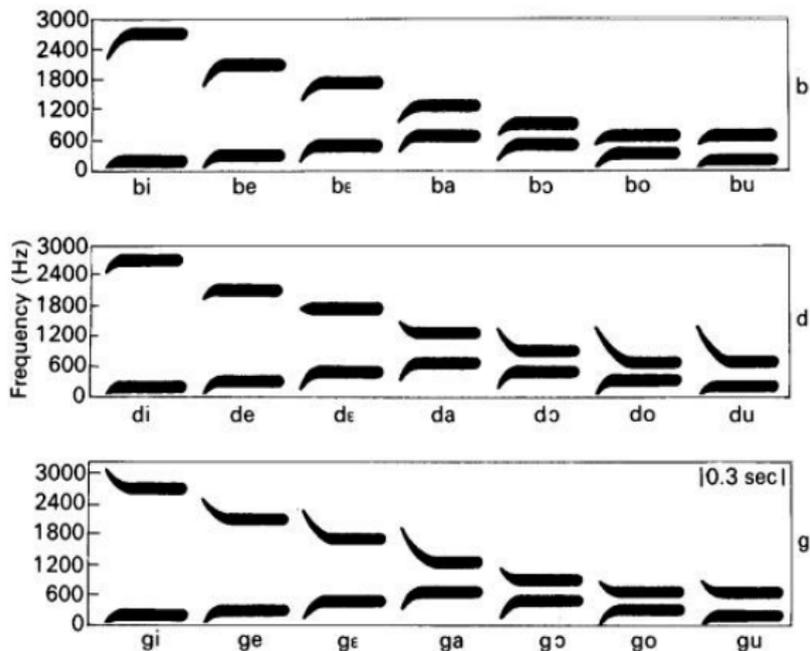


Aunque observamos duración de consonantes cortas y largas, observamos esta tendencia de relación de intercambio de duración.

El filtro oral y las oclusivas

- Los movimientos de cerrar y abrir ocurren en las vocales. Entonces las vocales llevan información importante del punto de articulación de las oclusivas.
- Las transiciones de formantes acercando y saliendo de las oclusivas indican su punto de articulación.
- Se puede estimar la trayectoria del formante saliendo de la oclusiva con una *frecuencia locus*. La transición se llama una *formula de locus* y involucra una aproximación polinomial a la trayectoria del formante.
- Las frecuencias de locus son 300 Hz. para bilabiales, 1800 Hz. para alveolares y más altos para velares/palatales (2200-2400 Hz.) (Delattre et al., 1955; Sussman et al., 1991).

El filtro de la cavidad oral y la transición C-V



(Delattre et al., 1955)

Las transiciones V-C

- Las transiciones de las vocales a las oclusivas (V-C) tienen propiedades diferentes de las transiciones C-V.
- La mayoría de los contrastes de punto de articulación se indica por la transición C-V pero unos se indica mejor por la transición V-C. Consonantes retroflejas se marca más por la transición V-C.
- Las consonantes velares se indica por un “pellizco” entre la ubicación de F2 y F3, pero depende de la calidad de la vocal que lo precede. Es más difícil observarlo con vocales posteriores.

Africadas

Se distinguen las africadas de las fricativas por la diferencia en el tiempo de ascenso de energía acústica (rise-time). Se mide la inclinación del cambio en amplitud al principio de la fricación.

Es más gradual con fricativas y abrupto con africadas.

(Ejemplo en Praat)

Ejercicio: Cuáles son los rasgos más importantes?

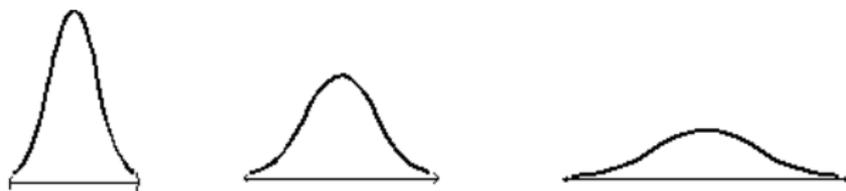
Meta: Qué pasa si cambiamos las rupturas?

- 1 Ponga la aspiración de [p] en [t, k]. Es [p] todavía?
- 2 Ponga la aspiración de [t] en [p, k]. Es [t] todavía?
- 3 Ponga la aspiración de [k] en [p, t]. Es [k] todavía?

La resonancia y el aflojamiento

Las resonancias en la cavidad oral aflojan las armónicas (teoría fuente-filtro). Igualmente otras articulaciones pueden aflojar las resonancias.

- Tejido suave afloja señales acústicas.
- Constricción excesiva afloja señales acústicas, p.ej. el amplitud muy bajo de F2 y F3 con [u].
- Ventilación del aire por otro canal aflojará el amplitud de las resonancias en el canal central, p.ej. los senos nasales.



Cuando se aflojan los formantes, se aumenta su ancho de banda.

El ancho de banda de formantes es el rango de frecuencias que está incluido en la cima espectral. Una cima más coronada de un formante tiene un ancho de banda estrecho. Un formante amplio tiene un ancho de banda más grande.

Por qué?

Nunca observamos las resonancias – solamente vemos su efecto en las armónicas. Desde que hay una constricción a través de una región cerca de un punto de articulación (y no en un punto específico), los formantes siempre reflejan un rango de frecuencias.

Nasales

En la producción de una oclusiva nasal como [N], el tubo está cerrado en un lado y casi siempre abierto en el otro.

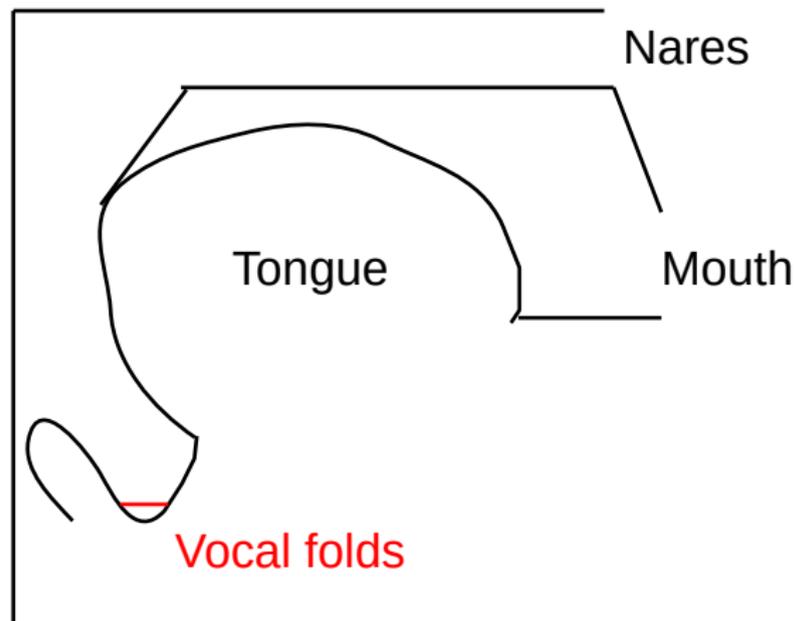
Hay resonancias en la cavidad nasal igualmente como en la cavidad oral. **Los formantes nasales** se calcula por la longitud del tubo (21.5 cm para muchas personas.)

La formula de resonancia para un tubo con un lado cerrado:

$$F_n = (2n - 1)c/4L.$$

El tejido suave en la cavidad nasal y la constricción crean formantes más aflojados con un ancho de banda más grande de lo que se ve con vocales.

Las cavidades faringales y nasales como un tubo uniforme en la producción de un nasal uvular. La distancia promedio entre la uvula y los nares es 12.5 cm.



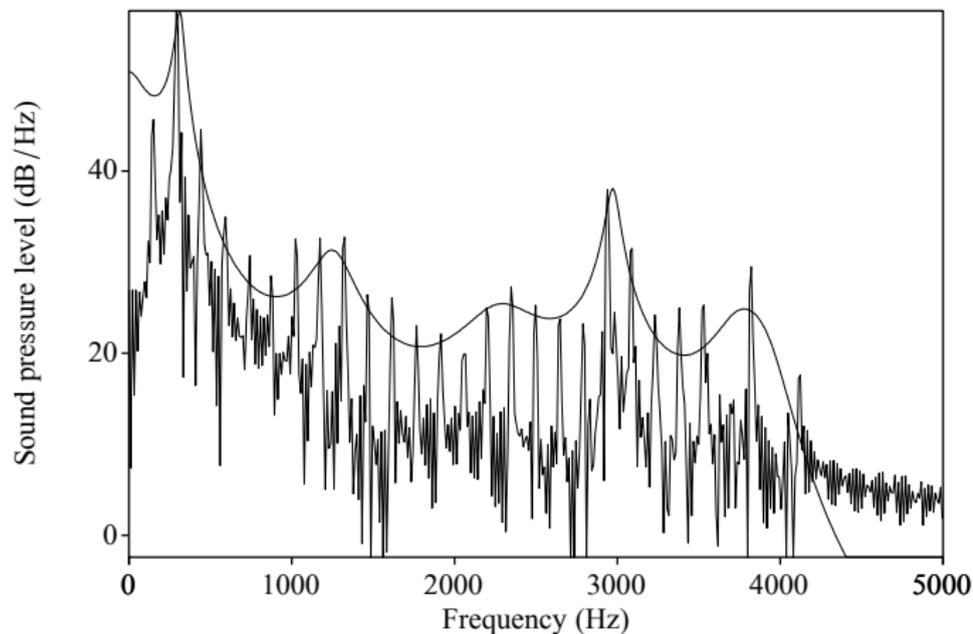
Podemos calcular las resonancias de este tubo de (21.5 cm) usando la formula $F_n = (2n - 1)c/4L$

Qué sería N1 para un tubo de 21.5 cm? N2? N3?

Generalmente esta formula funciona bien pero los nares siempre producen una constricción más estrecha que la cavidad nasal. Entonces esta constricción baja la frecuencia de los formantes nasales (igualmente como los labios en la cavidad oral).

Predecimos un formante nasal cada 800 Hz.

El espectro de [ŋ]



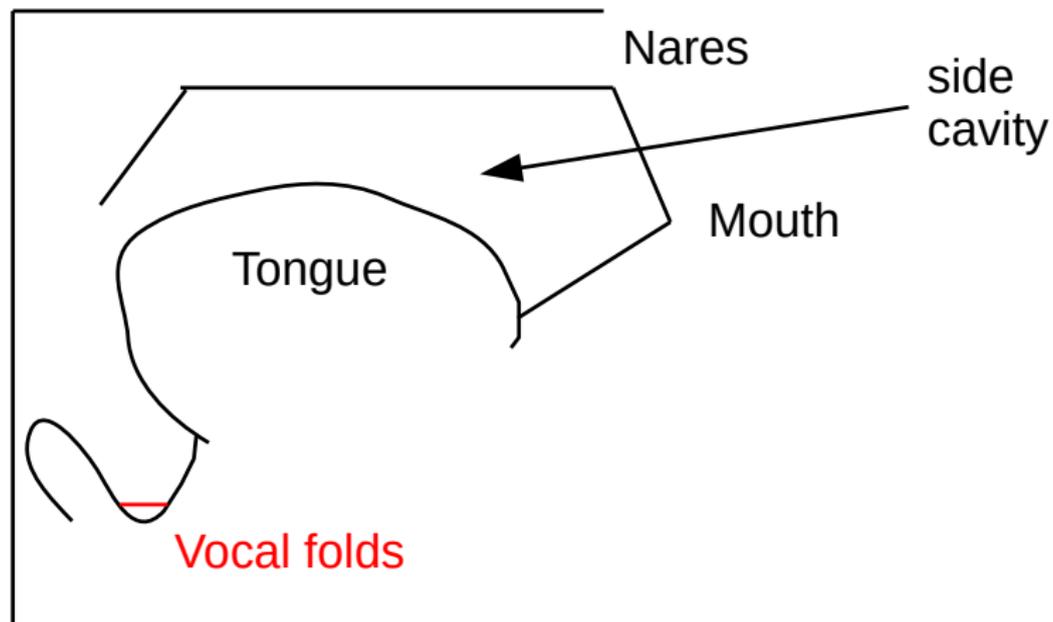
Las cavidades nasales

Los formantes nasales tienen un amplitud más bajo de formantes orales por las constricciones adicionales de la cavidad nasal y las cavidades de senos.

Cuando producimos consonantes nasales en puntos de articulación más anteriores, producimos una cavidad al lado de la cavidad oral.

Estas cavidades también tienen resonancias! Podemos calcularlas como si fueran un tubo uniforme con un lado cerrado.

La producción de un nasal bilabial incluye una cavidad al lado de aproximadamente 8 cm.



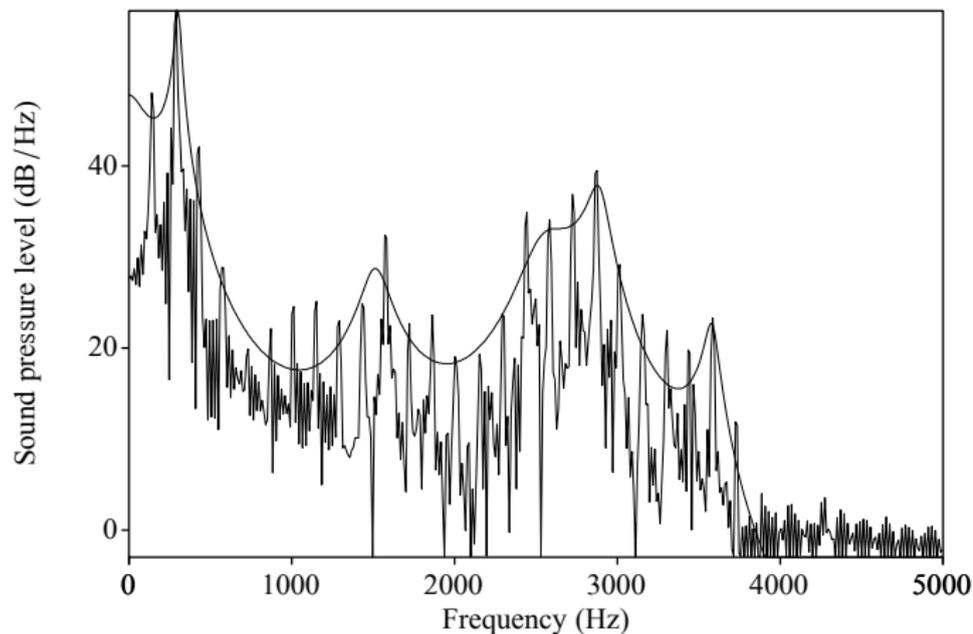
Anti-formantes

Las resonancias de la cavidad al lado no amplifican las resonancias de la señal, sino producen interferencia destructiva.

Las frecuencias de resonancia de la cavidad del lado que están cerca de las resonancias del tubo nasal se disminuyen. Las resonancias de la cavidad del lado se llaman **anti-formantes**.

Anti-formantes crean valles en amplitud en el espectro, no cimas.

El espectro de [m]

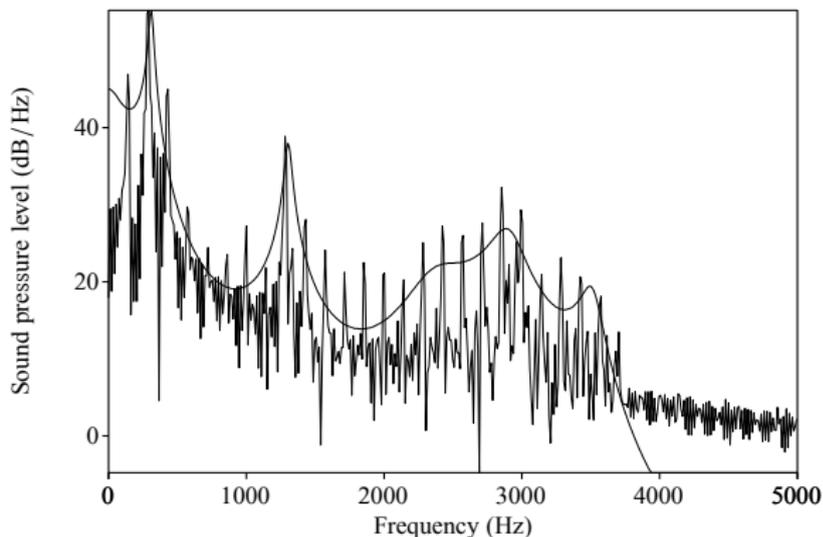


La primera resonancia de la cavidad del lado (de 8 cm) ocurre en 1094 Hz. Es aproximadamente la ubicación del valle en el espectro de [m].

Adicionalmente, los anti-formantes disminuyen las resonancias altas también. Podemos observar esto en el espectro de [n] (la próxima diapositiva).

El espectro de [n]

El primer anti-formante de la cavidad al lado (de 5 cm) ocurre en 1750 Hz. Observe donde se ubica la valle.



Los senos también producen anti-formantes en la producción de la nasalización pero su morfología varía a través de personas diferentes. Entonces no son muy confiables.

Más de todo confiamos en los formantes nasales en la percepción de consonantes nasales. Tienen amplitudes más bajas y un ancho de banda más grande de los formantes vocálicos.

También usamos los índices transicionales de C-V, V-C para predecir el punto de articulación de la oclusiva nasal (las formulas locus).

Ejemplos de nasales diferentes con vocales y sin vocales adyacentes en Praat.

Aproximantes laterales

Se baja los lados de la lengua en la producción de aproximantes laterales.

Crea una cavidad al lado que produce un anti-formante. Para un alveolar lateral, la cavidad es corta - 4 cm - y el antiformalte será de frecuencia alta.

El espaciado de los formantes es más ancho con aproximantes laterales que nasales porque la cavidad oral es más corta. Los formantes tienen amplitudes más altas también.

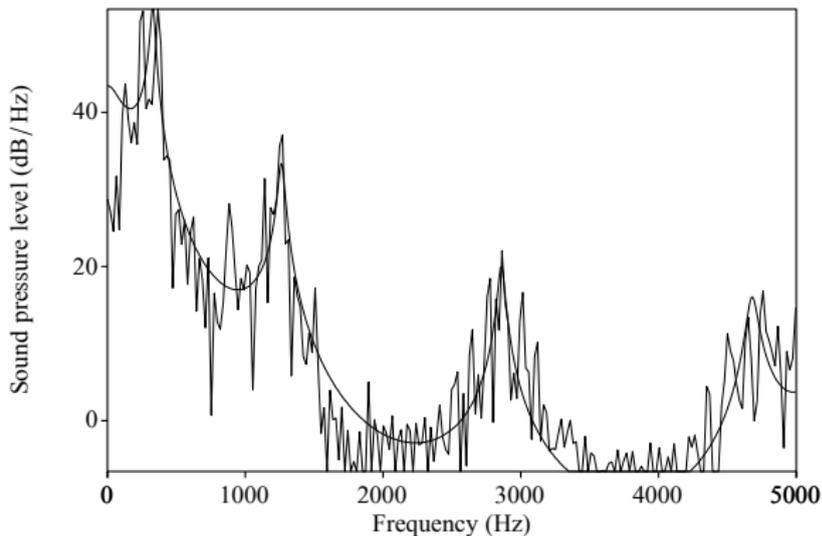
Si la cavidad del lado es más grande, tiene resonancias más bajas.

Laterales anteriores (apico-alveolar, lamino-dental) tienen resonancias bajas, un F2 entre 1200 - 1400 Hz.

Laterales posteriores (post-alveolar, palatal) tienen resonancias altas, un F2 entre 1600 - 2000 Hz.

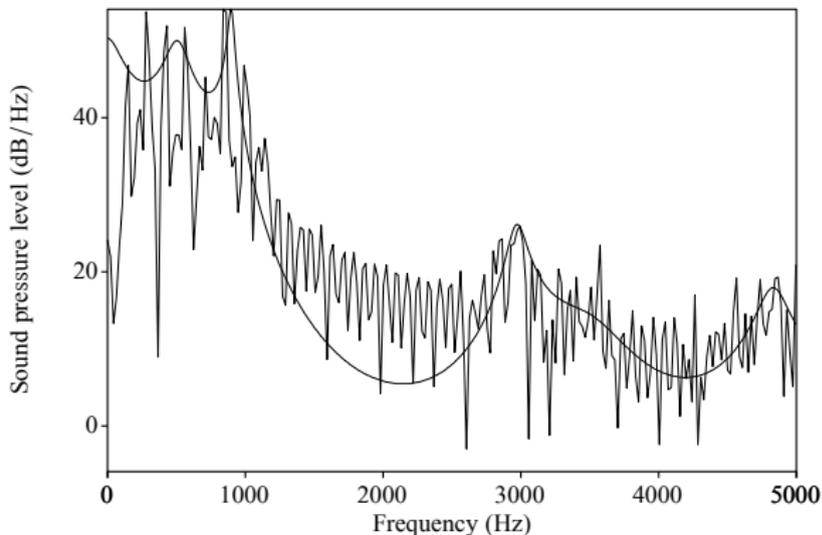
El espectro de [l] en 'a leaf'

Observe el espaciado de los formantes, sus amplitudes, y la ubicación de los valles. Con una cavidad del lado de 4 cm, el primer anti-formante estaría en 2187 Hz. Observe donde se ubica esta valle.



El espectro de [t] en 'tell him'

Observe lo que parece mucho al espectro de [u] (F1 y F2 muy bajo por la velarización), pero tiene menos amplitud y un aflojamiento de los formantes altos. Casi no hay una cavidad del lado y poca evidencia de anti-formantes.



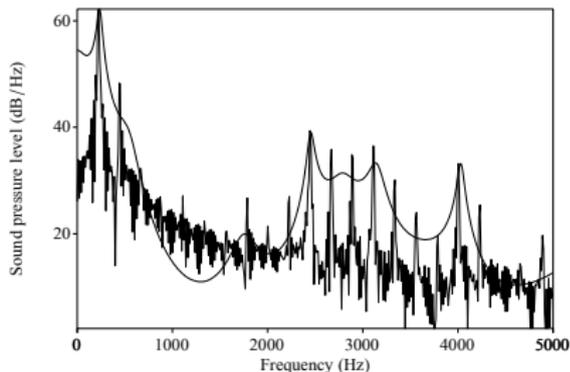
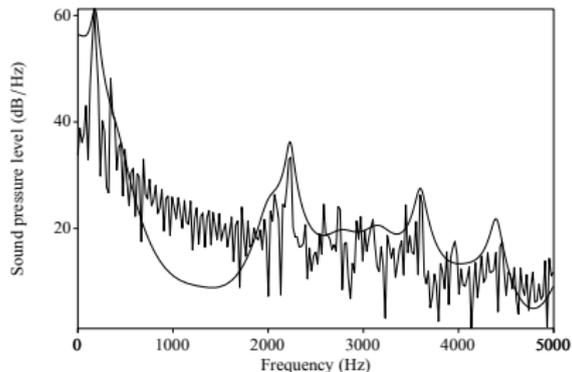
La nasalización vocálica

Vocales nasales son unos de los sonidos más complejos acústicamente. Por la presencia de las cavidades del lado, hay resonancias nasales y orales.

Sin embargo, las resonancias también son anti-formantes que producen interferencia destructiva contra las resonancias de la cavidad oral.

La fuerza de los anti-formantes depende del grado del enganche acústico entre las cavidades orales y nasales. Con un enganche débil, los antifonantes serán un poco más altos de los formantes nasales. Con un enganche fuerte, podemos calcular la frecuencia de los anti-formantes como una resonancia de la cavidad del lado, p.ej. un tubo de 12.5 cm de la uvula a las nares.

Las vocales de Triqui de Itunyoso [i, e, a, o, u, ï, ã, ã] (DiCanio, 2008, 2010). Observe el espectro de la vocal oral [i] (a la izquierda) en la palabra [ttʃiʔ²] 'diez' y la vocal nasal [ĩ] (a la derecha) en la palabra [stĩ⁴] 'uña'.

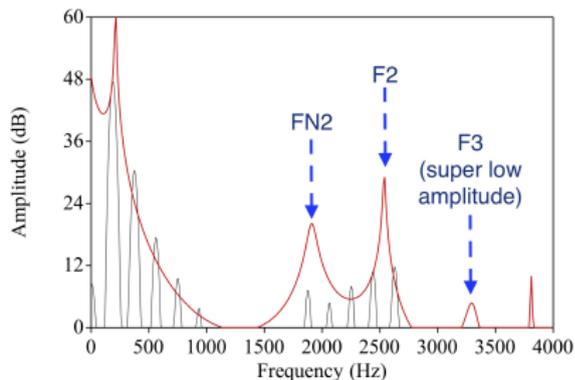
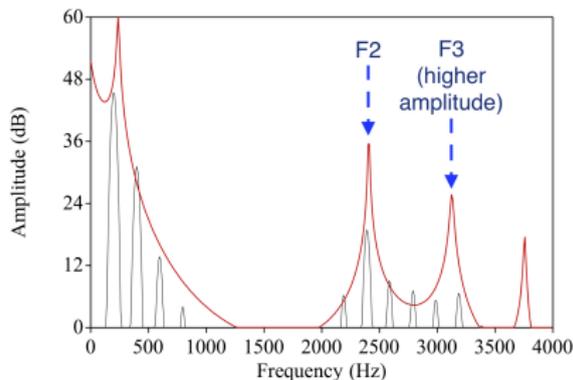


Hay estructura adicional en los formantes a la derecha.

Qué busco?

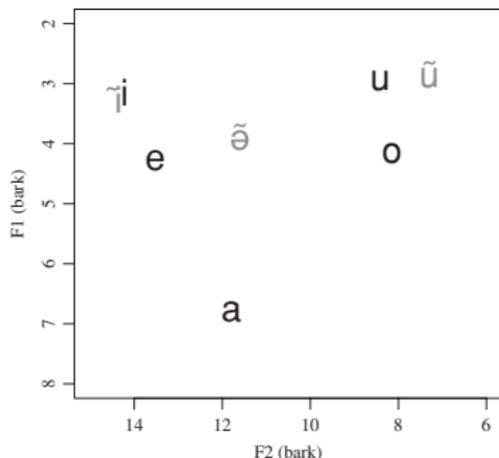
El espectro de [i] está a la izquierda y el espectro de [ĩ] está a la derecha.

Las diferencias acústicas pueden ser sutiles. **Siempre** se debe analizar vocales nasales al respecto de las vocales orales.



La nasalización vocálica

Desde que hay anti-formantes en el rango alto de valores de F1 (800 - 1000 Hz), tiene el efecto de elevar la altura de la vocal en la percepción. Vocales nasales suenan más altas que vocales orales.



- Cho, T. and Ladefoged, P. (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics*, 27:207–229.
- Delattre, P. C., Liberman, A. M., and Cooper, F. S. (1955). Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America*, 27(4):769–773.
- DiCano, C. T. (2008). *The Phonetics and Phonology of San Martín Itunyoso Trique*. PhD thesis, University of California, Berkeley.
- DiCano, C. T. (2010). Illustrations of the IPA: San Martín Itunyoso Trique. *Journal of the International Phonetic Association*, 40(2):227–238.
- DiCano, C. T. (2012). The Phonetics of Fortis and Lenis Consonants in Itunyoso Trique. *International Journal of American Linguistics*, 78(2):239–272.
- Johnson, K. (2012). *Acoustic and Auditory Phonetics*. Wiley-Blackwell, 3rd edition.
- Johnson, K., DiCano, C. T., and MacKenzie, L. (2007). The acoustic and visual phonetic basis of place of articulation in exrescent nasals. *Annual Report of the Phonology Laboratory at the University of California, Berkeley*, 3:529–561.
- Maddieson, I. (1997). Phonetic Universals. In Hardcastle, W. J. and Laver, J., editors, *The Handbook of Phonetic Sciences*, chapter 20, pages 619–639. Blackwell Publishing.
- Stevens, K. N. (2000). *Acoustic Phonetics*. MIT Press, first edition.
- Sussman, H. M., McCaffrey, H. A., and Matthews, S. A. (1991). An investigation of locus equations as a source of relational invariance for stop place categorization. *Journal of the Acoustical Society of America*, 90(3):1309–1325.