

Desmistificando la Fonética Acústica

*Mariano Pablo Mazzeo*¹

Introducción

La fonética acústica es tanto una rama de la lingüística como lo es de la física. Sin embargo, basta una rápida revisión bibliográfica para darse cuenta de que el lugar que ocupa en la lingüística es, en el mejor de los casos, marginal. Esto se debe en gran medida al grado de dificultad (en parte real y en parte sólo aparente) que representa para el docente/lingüista iniciarse en esta disciplina. A lo largo de este trabajo analizaremos e intentaremos arrojar luz sobre los principales puntos de conflicto para ayudar a romper la barrera que mantiene a los docentes/lingüistas del lado de afuera de esta disciplina.

Esas imposibles ecuaciones

La imagen 1 (ver apéndice de imágenes) fue subida a cuatro grupos de facebook de docentes y futuros docentes con el fin de sondear sus reacciones. Los resultados (apéndice I) fueron, como era de esperar, negativos. La mayor parte de los participantes expresaron, no sólo desconocimiento de estas ecuaciones, sino también dudas acerca de su propia capacidad de resolverlas.

¹ Profesor de Inglés en el IFS Pbro. Dr. Antonio María Sáenz, realizó la adscripción a la materia *Lengua y Expresión Oral I* y es Diplomado en la carrera de Diplomatura en Pedagogía de la Fonética con orientación en Lengua Inglesa de la UNSAM. En la actualidad, se encuentra cursando la carrera de Licenciatura en Inglés de la UNLIT. Licenciatura en Inglés – UNLIT.

La verdad es que no sólo estas ecuaciones son bastante más fáciles de resolver de lo que a primera vista se deduce², sino que ellas no ocupan lugar alguno en el uso que el docente/lingüista va a hacer de la fonética acústica; no hay necesidad alguna de que el profesional sepa resolverlas.

Circuitos electrónicos equivalentes

Muy comunes en la literatura sobre el tema son los ejemplos que se valen de circuitos electrónicos equivalentes³; particularmente a la hora de explicar los conceptos de “impedancia”, “resistencia” y “reactancia” acústicas. Una vez más, no sólo no es necesario para el docente/lingüista que utiliza la fonética acústica entender dichos circuitos, sino que tampoco le es necesario entender tales conceptos.

La onda sonora

Según Wikipedia, una onda sonora es “una onda longitudinal que transmite lo que se asocia con sonido. Si se propaga en un medio elástico y continuo genera una variación local de presión o densidad, que se transmite en forma de onda esférica periódica o cuasiperiódica”. Para entender lo que esto significa, nos vamos a valer del ejemplo de una cuerda (imagen 2). Como podemos ver, la cuerda está rodeada por partículas de aire (a). Al pulsarla, el desplazamiento de la cuerda va a desplazar a sus partículas vecinas (b) y al volver, (c), el vacío generado entre ésta y las partículas (rarefacción) va a generar un efecto de succión, atrayéndolas una vez más hacia la cuerda. Este proceso, a su vez, se va a repetir entre las partículas de aire adyacentes a la cuerda y sus partículas vecinas, propagando la onda sonora.

La percepción de una onda sonora depende de que su propagación alcance nuestro tímpano, replicándose en éste.

El senoide

Si bien la senoide (imagen 3) tampoco en sí mismo va a formar parte de las prácticas del docente/lingüista, es necesario entender qué representa, como base para entender aquellos conceptos que sí serán parte de su práctica docente.

² El lector interesado en saber cómo resolverlas encontrará vínculos a este fin en la sección de referencias.

³ El lector interesado también encontrará vínculos explicativos acerca de este tema en la sección de referencias.

La sinusoides (en fonética acústica) representa el movimiento de las partículas de aire. Lo que no es entendible a simple vista, sin embargo, es que la dimensión horizontal (izquierda-derecha) de la imagen no es una dimensión espacial sino temporal. En una experiencia realizada entre 38 alumnos de un curso de profesorado de inglés (apéndice II), en la que les fue presentada la onda sonora sin aclarar este punto, sólo 3 alumnos entendieron el concepto sin explicación previa.

Tal vez la mejor manera de interpretar cómo funciona la sinusoides es la de apelar a la analogía del video. Un video es una serie de imágenes estáticas de un mismo objeto en diferentes puntos en el tiempo. De la misma manera, aunque la partícula de aire se mueve en sólo dos direcciones (imagen 4), la sinusoides puede ser imaginada como una serie de imágenes estáticas de una partícula en diferentes puntos en el tiempo (imagen 5).

La correcta comprensión de este concepto nos facilita enormemente a la vez la tarea de comprender la información que la sinusoides nos brinda acerca de la onda sonora, a saber, frecuencia y amplitud.

Frecuencia

Según el diccionario de la Real Academia Española, la frecuencia es el “número de veces que se repite un proceso periódico por unidad de tiempo”.

En el caso de la onda sonora, la frecuencia está dada por la cantidad de veces que se completa un ciclo (en el caso de la sinusoides, el espacio entre un valor y su repetición luego de alcanzarse los valores máximos y mínimos) en un determinado período de tiempo (normalmente 1 segundo). En la imagen 6, por ejemplo, podemos ver que cada ciclo dura aproximadamente una décima de segundo. En otras palabras, la frecuencia es de alrededor de 10 ciclos por segundo, o 10 hertz.

El “tono” es el correlato psicoacústico de la frecuencia: es más grave cuando la frecuencia es más baja y más agudo, cuando es más alta.

Amplitud

En la sinusoides, la amplitud está definida por los valores máximos y mínimos (nótese su simetría), que equivalen en la onda sonora a los momentos de mayor desplazamiento de la partícula de aire desde su ubicación original. El correlato psicoacústico de la amplitud es el volumen⁴.

⁴ Si bien, dada esta definición, amplitud y volumen se entienden como conceptos diferentes, la idea de volumen como cantidad de aire desplazado es, por ser intuitiva, didácticamente útil.

En la imagen 7, podemos ver una onda sonora cuyos puntos máximos y mínimo corresponden a aproximadamente 1 decibelio⁵ y -1 decibelio, en otras palabras, su amplitud es de alrededor de 1 decibelio⁶.

Armónicos

Está claro que no es posible explicar la complejidad de los sonidos humanos valiéndonos solamente de la sinusoide. Acústicamente, los segmentos fonéticos son ondas sonoras complejas, compuestas de todo un espectro de sinusoides a las que llamamos “armónicos”. Los armónicos guardan entre sí una relación matemática, formando una serie en la que la frecuencia de cada nuevo armónico está dada por el producto de su número de orden y la frecuencia del primer armónico (también llamado frecuencia fundamental [f₀]). Por ejemplo, si la frecuencia fundamental es de 100hz, la del segundo armónico va a ser de 200hz (2x100), la del tercero de 300hz (3x100), la del cuarto de 400hz (4x100), y así *ad infinitum*⁷.

Dado que los algoritmos utilizados para calcular la frecuencia fundamental pueden generar errores, el hecho de que los armónicos sean múltiplos de ésta, y acompañen por lo tanto sus contornos, es particularmente útil en los estudios de entonación. La imagen 8 corresponde a un espectrograma de banda angosta (el tipo que permite ver los armónicos) de un tono ascendente. En él, podemos apreciar como todos los armónicos (en negro) ascienden junto con la frecuencia fundamental (en azul).

Resonancia

La resonancia es la respuesta vibratoria de un cuerpo a un estímulo. En otras palabras, si pulso una cuerda, vibra.

⁵ La escala de decibelios es una escala logarítmica en la que la amplitud de un sonido se compara con un valor de referencia (en general 20μP).

⁶ Debido a la simetría de toda onda periódica, no es posible promediar sus valores en la manera tradicional, y se utiliza en cambio el sistema RMS (*root-mean-square*) en el que los valores son primero elevados al cuadrado, luego promediados, y el resultado es la raíz cuadrada del promedio; en este caso $\sqrt{((1+12)/2)}=1$.

⁷ Si bien en teoría no existe un límite máximo para el número de armónicos, en la práctica, el espectro pierde alrededor de 6dB por cada octava, por lo que es difícil encontrar armónicos por encima de la quinta octava.

Todo cuerpo vibra con una frecuencia específica, y esta regla no sólo se aplica a los sólidos: los líquidos y gases contenidos en un recipiente o cavidad también experimentan este fenómeno. Lo que puede deducirse de este dato es que el aire en las cavidades del tracto vocal va a resonar a frecuencias específicas, determinadas por la forma y dimensiones que éste asuma para cada segmento. El efecto práctico de este fenómeno es que los armónicos cuyas frecuencias se aproximen a las frecuencias de resonancia del tracto vocal van a verse amplificados.

Para entender por qué se produce este efecto, imaginemos a un atleta saltando la soga. Si la frecuencia con que la soga realiza un giro completo es diferente a la frecuencia con que el atleta salta, la soga va a ocasionalmente impactar los tobillos del atleta e interrumpirá la fluidez del movimiento. Si la frecuencia, en cambio, es la misma, la soga se va a mover con libertad y alcanzará su máxima velocidad. De la misma manera, las partículas de aire van a moverse con más libertad a medida que se aproximen a la frecuencia de resonancia del tracto vocal y transmitirán así el movimiento a una mayor cantidad de otras partículas.

Formantes

La pieza faltante del rompecabezas acústico son los formantes. Los formantes son picos espectrales resultantes de la resonancia del tracto vocal. En otras palabras, son las frecuencias alrededor de las cuales la resonancia del tracto vocal amplifica los armónicos⁸. Para poner un ejemplo práctico, el primer formante de una vocal abierta (digamos |a|) se encuentra alrededor de los 700hz, lo que significa que para una frecuencia fundamental de 80hz, el octavo y noveno armónicos (640 y 720hz respectivamente) van a estar más amplificados que los anteriores.

¿Qué información podemos obtener de los formantes?

Más importante que saber de dónde vienen es saber qué información nos proporcionan los formantes acerca de los segmentos que los producen.

⁸ Como veremos a continuación, las frecuencias que no forman parte de una secuencia de armónicos también son amplificadas alrededor de estas frecuencias.

Vocales

En el caso de las vocales, son los dos primeros formantes los que nos brindan la mayor parte de la información. En la imagen 9, podemos ver un espectrograma de banda ancha (el tipo que permite ver los formantes) de la secuencia [aeiou].

Como podemos observar, el primer formante (F1) más alto (alrededor de 650hz) corresponde a la vocal abierta ([a]), mientras que el más bajo (alrededor de 315hz) corresponde a las vocales cerradas ([i] y [u]). Esto nos indica que la altura del primer formante es inversamente proporcional a la altura de la lengua.

Por otra parte, también podemos observar que el segundo formante (F2) más alto (alrededor de 2250hz) corresponde a la vocal más anterior ([i]), mientras que el más bajo (alrededor de 870hz) corresponde a las vocales posteriores ([o] y [u]). Esto nos indica que la altura del segundo formante es proporcional a la anterioridad de la lengua.

La relación entre los formantes y las posiciones de la lengua se ve reflejada en la imagen 10.

¿Y qué hay de los labios?

No existe total consenso acerca de qué formantes son afectados por la posición de los labios, con algunos autores (por ejemplo, Harrington, 1999, p. 58) que afirman que afecta al segundo formante y otros (por ejemplo, Wood, 2014, p. 1), que afecta a todos. En cualquier caso, dado que el segundo formante es el de mayor rango, es entendible que el efecto sea más notorio en este último. Sí existe consenso, sin embargo, acerca de la manera en que los afecta, con la posición extendida aumentando las frecuencias y la posición redondeada disminuyéndolas. La imagen 11 corresponde a un espectrograma de la secuencia [iy], en el que puede observarse este efecto.

Consonantes

En el caso de las consonantes, necesitamos incorporar el tercer formante, así como también los conceptos de aperiodicidad y antiformante.

La importancia del segundo y tercer formantes

Los movimientos del segundo y tercer formantes en las transiciones hacia y desde los sonidos oclusivos son un dato clave para identificar su punto de

articulación. La imagen 12, por ejemplo, corresponde a un espectrograma de las secuencias |ebel|, |ada| y |aga|. En este podemos observar como las transiciones entre las vocales y los oclusivos están marcadas por un descenso de los formantes para la bilabial ([b]), un ascenso para la alveolar ([d]) y un encuentro (“*velar pinch*”) para la velar ([g]).

El tercer formante es también de particular importancia en la identificación de los sonidos laterales, para los que es bastante alto (alrededor de 3000hz), y de algunos aproximantes, por ejemplo [ɹ], para la que es particularmente bajo (menor a 2000hz).

Ondas aperiódicas

El nombre *onda aperiódica* es el término técnico con el que nos referimos al ruido. El ruido se caracteriza por no tener una frecuencia fundamental o una secuencia de armónicos definidas. El ruido está compuesto, en cambio, por frecuencias erráticas.

Esto no significa, sin embargo, que el ruido no aporte información segmental. La imagen 13 corresponde a un espectrograma de las secuencias |fɔ:t|, |θɔ:t|, |sɔ:t| y |ʃɔ:t|. La primera distinción clara es que los sibilantes tienen más energía que los no sibilantes. A su vez, podemos distinguir entre estos porque para la alveolar ([s]) la mayor concentración de energía comienza alrededor de los 4000hz, mientras que para la postalveolar ([ʃ]), comienza alrededor de los 2500hz.

Con respecto a los no sibilantes, la distinción es sutil (como lo es en la percepción), con más presencia de volumen en las frecuencias bajas para la labio-dental ([f]).

¿Y qué hay de la |h|?

Las frecuencias presentes en una onda no periódica también pueden generar resonancia en el tracto vocal. Por esta razón, el segmento comúnmente representado como |h| presenta en el espectrograma formantes muy similares a los de la vocal que le sigue, aunque más tenues⁹. La imagen 14 corresponde a un espectrograma de las secuencias |hɔ:d|, |hed| y |hid|, en las que puede observarse este efecto.

⁹ Por esta razón, algunos autores (por ejemplo, Laver, 1994, pp. 304-5) sugieren que este segmento puede ser clasificado como una versión sorda de dicha vocal.

Combinación de fuentes periódicas y aperiódicas

Los obstruyentes sonoros combinan elementos periódicos con elementos aperiódicos. La imagen 15 corresponde al espectrograma de la consonante [z]. Como se puede observar, la concentración de energía por encima de los 4000hz (presente en su contraparte sorda [s]) coexiste con la presencia de una serie de armónicos.

Antiformantes

Como puede deducirse de su nombre, los antiformantes son frecuencias alrededor de las cuales la resonancia (o más específicamente “antiresonancia”) del tracto vocal atenúa el volumen de los armónicos.

Los ejemplos más típicos de antiformantes se encuentran en los segmentos nasales, para los que la frecuencia del antiformante principal (“*nasal pole*”) puede ser interpretada como correlato del punto de articulación. La imagen 16 corresponde a un espectrograma de las secuencias |'sɪmə|, |'sɪnə| y |'sɪŋə|. Como puede observarse, el principal antiformante se encuentra alrededor de 800hz para la bilabial ([m]), alrededor de 950hz para la alveolar ([n]) y alrededor de 2400hz para la velar ([ŋ]).

Conclusión

Si bien es cierto que algunos de los elementos de la fonética acústica que mayor rechazo generan en el docente/lingüista no son necesarios en la práctica docente, hay otros que sí lo son. Tal es el caso de los conceptos de onda sonora, sinusoidal, armónico, resonancia, formante, antiformante y onda aperiódica. Sobre ellos, hemos intentado arrojar luz en este trabajo.

Apéndice de imágenes

Imagen 1. Ecuaciones más comúnmente usadas en fonética acústica

a)
$$X_k = \sum_{n=0}^{N-1} x_n \cdot e^{-i2\pi kn/N}$$

b)
$$e^{i\theta} = \cos\theta + i\sin\theta$$

c)
$$-\frac{\rho c}{A_1} \cot(kl_1) + \frac{\rho c}{A_2} \tan(kl_2) = 0$$

d)
$$z = \sqrt{R^2 + X^2}$$

$$k = \frac{2\pi f}{c} \Rightarrow f = \frac{kc}{2\pi}$$

a) Transformación de Fourier, b) Fórmula de Euler, c) Teoría de la perturbación, y d) Impedancia

Imagen 2. Movimiento de una cuerda y sus partículas de aire vecinas

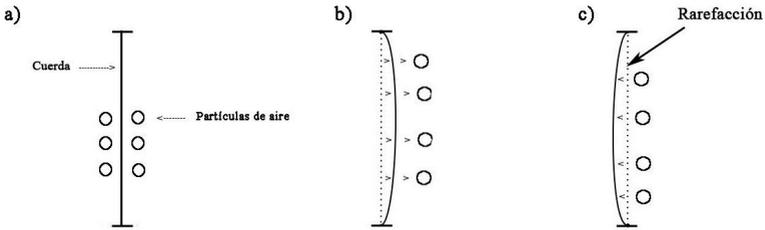


Imagen 3. Sinusoide

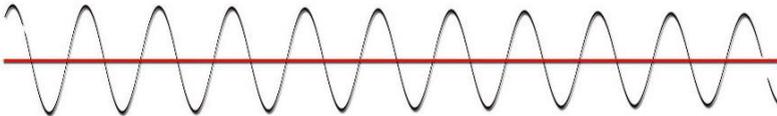


Imagen 4. Movimiento de la partícula de aire

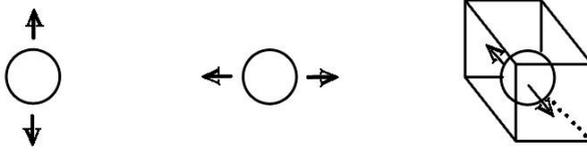


Imagen 5. Partículas de aire en la senoide

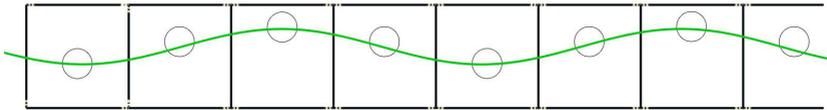


Imagen 6. Ejemplo de ciclo sinusoidal

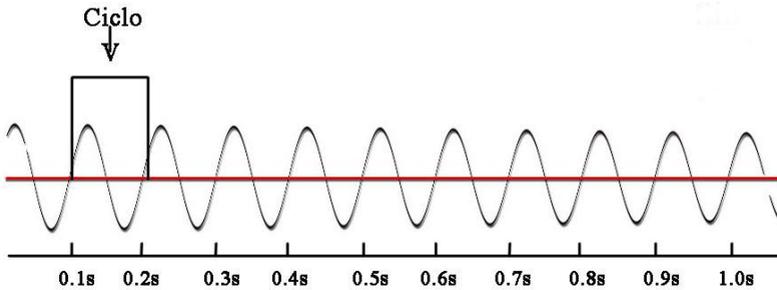


Imagen 7. Ejemplo de amplitud representada en el senoide

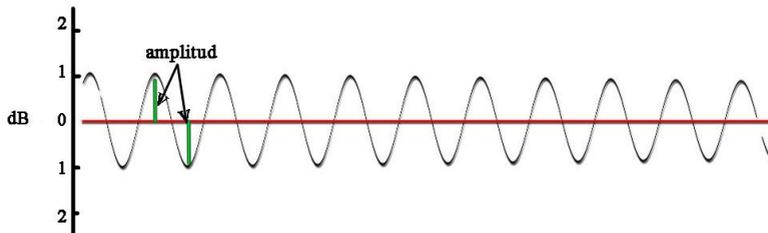


Imagen 8. Armónicos representados en el espectrograma

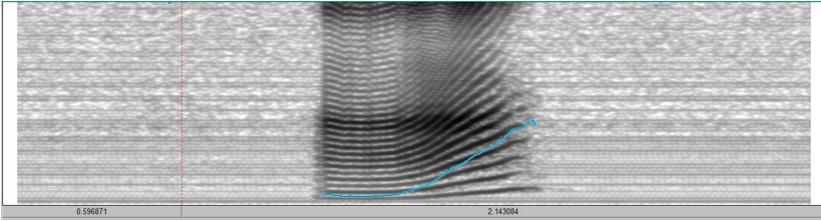


Imagen 9. Formantes representados en el espectrograma

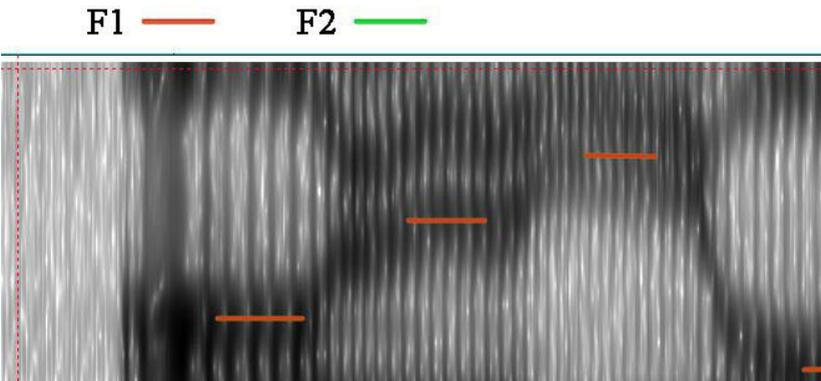


Imagen 10. Formantes representados en el cuadrilátero vocálico

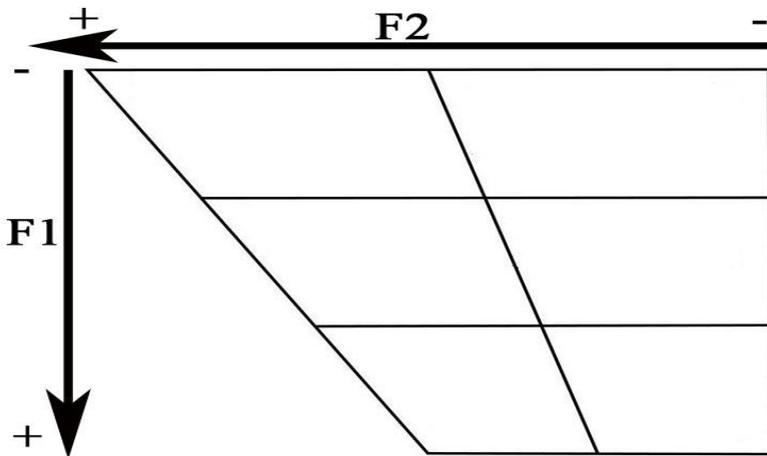


Imagen 11. Efecto de las posiciones labiales en los formantes

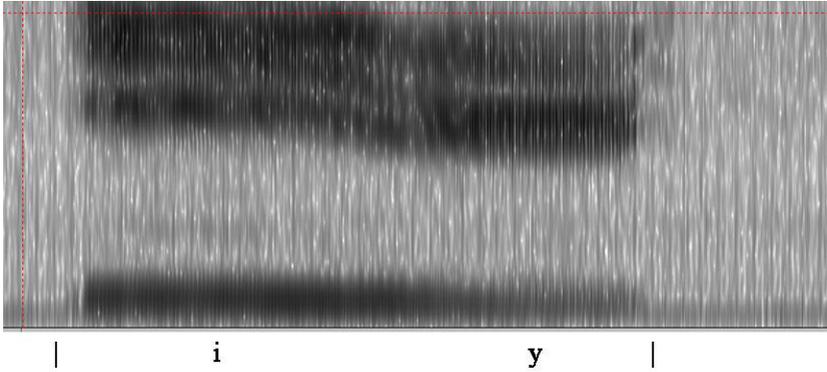


Imagen 12. Transiciones de los formantes en la vecindad de los segmentos oclusivos

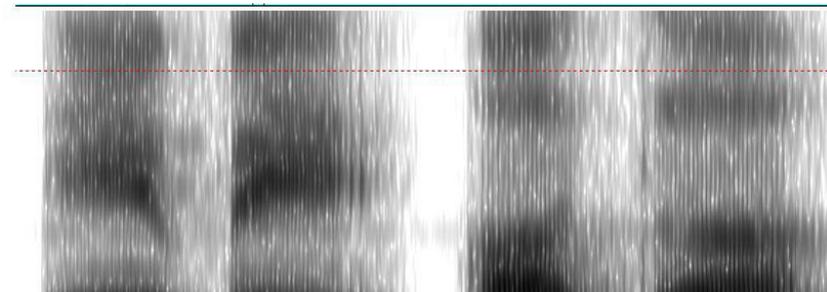


Imagen 13. Segmentos fricativos representados en el espectrograma

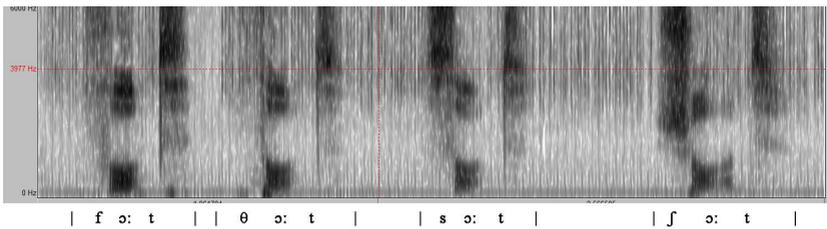


Imagen 14. Segmento [h] representado en el espectrograma

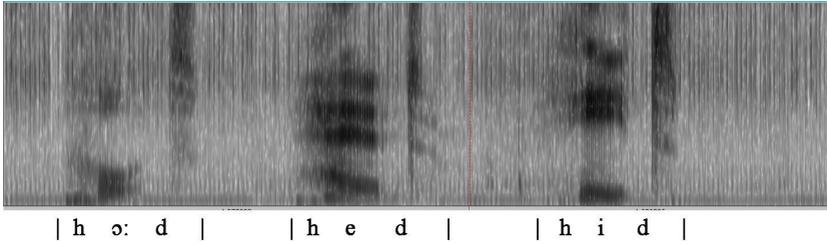


Imagen 15. Segmento [z] representado en el espectrograma

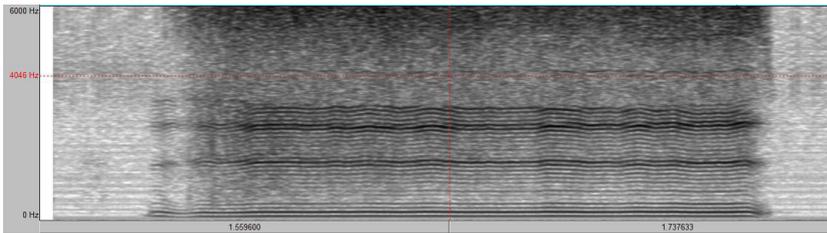
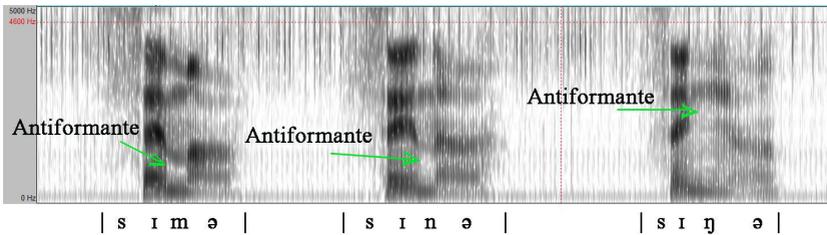


Imagen 16. Segmentos nasales representados en el espectrograma



Apéndice I

Respuestas negativas:

Frustración.

Something I would never be able to solve.

Lo haría mal seguramente.

Thanks God I'm a teacher of English.

It reminds me of school days but I cannot solve any of them.

Something impossible to solve!

Clueless.

¿A dónde se han ido todos los números?

Some sort of math! Ouch.

Algo que nunca, creo, poder aplicar en tiendas de comida.

Otro idioma.

A headache.

I don't like Maths.

Confusión.

Something impossible for me to do.

When I see this picture... I remember why I'm studying English!

Pain.

Headache.

When I see this I have a feeling of running away and ask the pass to kinder!!

Hahaha sorry but I'm not good at maths.

And then Satan said: "put the alphabet in math".

I feel confused.

It makes me wanna cry.

Respuestas positivas:

The beauty of Mathematics. The relationship that exists between Maths and Language.

Formulas created by Fourier, Euler, Pitagoras and others

Apéndice II

Total de alumnos: 38

Respuestas incorrectas: 21

Respuestas correctas: 3

No sabe/no contesta: 14

Referencias bibliográficas

Real Academia Española. (2014). Disquisición. En *Diccionario de la lengua española* (22.^a ed.). Recuperado de <http://dle.rae.es/srv/search/search?w=frecuencia>.

Glass, J. (2003). *Acoustic Theory of Speech Production. (Lecture 2 Handouts)*. Massachusetts Institute of Technology. Recuperado de <http://ocw.mit.edu/>

- [courses/electrical-engineering-and-computer-science/6-345-automatic-speech-recognition-spring-2003/lecturenotes/lecture2.pdf](#)
- Hagiwara, R. (s/f). *How to Read a Spectrogram*. Winnipeg: University of Manitoba. Recuperado de <http://home.cc.umanitoba.ca/~robh/howto.html>
- Harrington, J. (1999). *Techniques in Speech Acoustics*. Springer-Science+Business Media, BV.
- Haughton, P. (2002). *Acoustics for Audiologists*. UK: Academic Press.
- Hawkins, S. (2006). *Acoustic Theory of Speech Production (Supplement to and extension of paper 3 lectures)*. Cambridge: University of Cambridge. Recuperado de http://www.ling.cam.ac.uk/li9/m4_0708_acoustictheoryspeechproduction_07-8.pdf
- Huckvale, M. (2008). *Make your own vowel resonators!* London: University College London. Recuperado de <http://www.phon.ucl.ac.uk/home/mark/vowels/>
- Johnson, K. (2012). *Acoustic and Auditory Phonetics, Third Edition*. UK: Wiley-Blackwell.
- Joutsenvirta, A. (s/f). *The Basics of Acoustics*. Helsinki: Sibelius Academy. Recuperado de <http://www2.siba.fi/akustiikka/?id=42&la=en>.
- Krasnow, B. (2015). How a horn amplifies sound (hint: Impedance matching). *Youtube*. Recuperado de <https://www.youtube.com/watch?v=vcSc16tnVqk>.
- Krasnow, B. (2011). Tutorial: Electrical impedance made easy - Part 1. *YouTube*. Recuperado de <https://www.youtube.com/watch?v=xyMH8wKK-Ag>.
- Krasnow, B. (2011). Tutorial: Electrical impedance made easy - Part 2. *YouTube*. Recuperado de <https://www.youtube.com/watch?v=tZBMfDvWF4U>.
- Ladefoged, P. (1996). *Elements of Acoustic Phonetics, Second Edition*. USA and UK: The University of Chicago Press.
- Ladefoged, P. (2001). *Vowels and Consonants. An Introduction to the Sounds of Languages*. USA, Blackwell Publishers Ltd.
- Laver, J. (1994). *Principles of Phonetics*. Ca,bridge: Cambridge University Press.
- Mannell, R. (s/f). *Speech Acoustics*. Sidney: MacQuarie University. Recuperado de <http://clas.mq.edu.au/speech/acoustics/frequency/resonance.html>.

- Mazzeo, M. (2015). Acoustic Phonetics for Language Teachers. *YouTube*. Recuperado de <https://www.youtube.com/watch?v=xD9Q5eEORNU>.
- Mazzeo, M. (2016). Acoustic Phonetics for Language Teachers II. *YouTube*. Recuperado de <https://www.youtube.com/watch?v=TA5kVv1XVo>.
- Mazzeo, M. (2016). Acoustic Phonetics for Language Teachers III. *YouTube*. Recuperado de <https://www.youtube.com/watch?v=nDFePahaOjc>.
- Mazzeo, M. (2016). Acoustic Phonetics for Language Teachers IV. *YouTube*. Recuperado de <https://www.youtube.com/watch?v=cprRIxGbHKI>.
- Mazzeo, M. (2016). Acoustic Phonetics for Language Teachers V. *YouTube*. Recuperado de <https://www.youtube.com/watch?v=nPUfczpcgVc>.
- Mazzeo, M. (2016). Acoustic Phonetics for Language Teachers VI. *YouTube*. Recuperado de <https://www.youtube.com/watch?v=OvTeTPDP32M>.
- Mazzeo, M. (2016). Acoustic Phonetics for Language Teachers VII. *YouTube*. Recuperado de <https://www.youtube.com/watch?v=E9yhWZi2fh4>.
- Onda Sonora. (2018). *Wikipedia*. Recuperado de https://es.wikipedia.org/wiki/Onda_sonora.
- Piché, J. y Nix, P. (1997). *Formant Table III: Formant values*. Chicago: University of Chicago. Recuperado de https://www.classes.cs.uchicago.edu/archive/2000/spring/CS295/Computing_Resources/Csound/CsManual3.48b1.HTML/Appendices/table3.html.
- Raphael, L., Borden, G. y Harris K. (2011). *Speech Science Primer, sixth edition*. USA: Wolters Kluwer (Health).
- Stevens, K. y House, A. (1955). Development of a Quantitative Description of Vowel Articulation. *The Journal of the Acoustical Society of America*, 27(3). Recuperado de <http://homepages.wmich.edu/~stasko/sppa601/readings/stevens&house.pdf>.
- Stevens, K. (2000). *Acoustic Phonetics*. UK: The MIT Press.
- Tizte, I. (2009). How Are Harmonics Produced at the Voice Source? *Journal of Singing*, 65(5), 575-576.
- Wolfe, J. (s/f). *Music Acoustics*. Sidney: The University of New South Wales. Recuperado de <http://newt.phys.unsw.edu.au/jw/basics.html>.
- Wood, S. (2014). The Acoustic Weaknesses of the Bell Vowel Model. *SWPhonetics*. Recuperado de <http://swphonetics.com/articulation/bell-vowel-model/bellvowelacousticweaknesses/>.

- Wood, S. (2012). The Bell Vowel Model. *SWPhonetics*. Recuperado de <http://swphonetics.com/articulation/bell-vowel-model/>.
- Xu, S. (2015). Discrete Fourier Transform - Simple Step by Step. *YouTube*. Recuperado de https://www.youtube.com/watch?v=mkGsMWi_j4Q.
- Xu, S. (2015). The FFT Algorithm - Simple Step by Step. Recuperado de <https://www.youtube.com/watch?v=htCj9exbGo0>.