

Tono y frecuencia fundamental

Propuesta metodológica para su estudio¹

Domingo Román Montes de Oca

Diana Muñoz Builes

Una de las propiedades de los sonidos del habla, junto con el timbre, la intensidad y la duración, es la altura tonal, que está presente en las vocales y en las consonantes sonoras. Esta propiedad permite que se perciba un sonido como alto o bajo en relación a otro. La sucesión de estos tonos (altos y bajos) produce la percepción de una "línea melódica" del enunciado. Por ejemplo, en las preguntas absolutas, lo habitual es que terminen con una elevación del tono y gracias a ese rasgo prosódico se puede identificar su valor pragmático. Así, la diferencia de sentido entre "vas a venir" y "¿vas a venir?" se produce y se interpreta a partir de la sucesión de tonos, especialmente en la última parte del enunciado.

En este capítulo se presenta una metodología fonética que puede ser aprovechada para el estudio científico de este tema. Se presentan también los conceptos y los problemas metodológicos de un estudio de esta naturaleza.

Problemas y definiciones en el análisis del tono

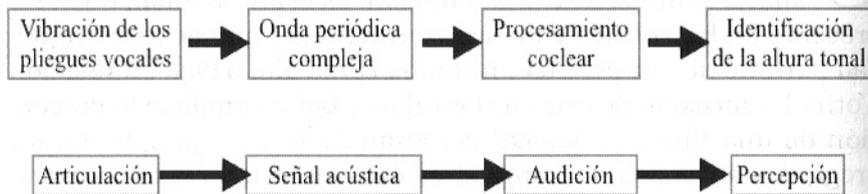
Para indagar fonéticamente en el tono de una emisión debemos partir de un valor acústico que se denomina frecuencia fundamen-

¹ Propuesta desarrollada para el proyecto Fondecyt 1130720 "Mapa prosódico de Chile. Descripción fonética y fonológica de rasgos suprasegmentales del español hablado por chilenos".

tal y se simboliza como f_0^2 . Este valor se obtiene de las señales periódicas del habla y, a su vez, estas señales son producto de la vibración de los pliegues vocales. Así, este fenómeno tiene una dimensión articulatoria o fisiológica (la vibración regular de los pliegues vocales) y un aspecto acústico (la periodicidad de la señal). Además, hay un momento de procesamiento auditivo, que resulta totalmente inconsciente, pues las señales son procesadas en la cóclea y transformadas en señal eléctrica para su procesamiento neural. Todo este proceso termina en la percepción consciente de esa sensación de más alto o más bajo, que es lo que se llama propiamente "tono". Este proceso se representa esquemáticamente en la figura 1.

Figura 1

Esquema de los pasos implicados desde la producción hasta la percepción del tono

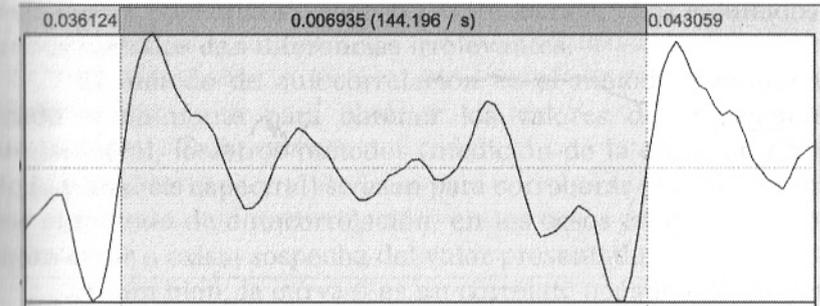


Los programas de análisis acústico como Praat (Boersma y Weenink, 2016), que se usará en este estudio, permiten analizar de manera manual y automática el valor de la frecuencia fundamental, y a partir de ese dato poder inferir cuestiones sobre el tono.

La forma para obtener manualmente el valor de la f_0 es acercarse a la señal hasta que se pueda ver en la ventana un ciclo y medirlo. Dado que la frecuencia es la inversa del período (T) y el período es la duración del ciclo, basta con conocer el valor de T y hacer la operación de dividir $1/T$. El programa Praat presenta el resultado de esa operación en la parte superior de la ventana de edición.

2 Para una descripción más profunda de los problemas teóricos y técnicos en la determinación del *pitch*, véase Chevgne (2006), Ladefoged (2003) y Johnson (2011).

Figura 2
Acercamiento a en un oscilograma que muestra la selección de un ciclo

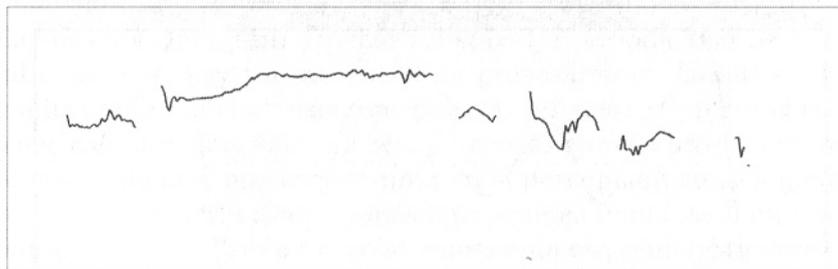


En la figura 2 se puede observar la imagen del oscilograma en la que se ha seleccionado un ciclo. En la parte superior de la selección se lee el valor de la duración (de la parte seleccionada) y en el paréntesis el valor de 1 dividido por esa duración.

Este método manual es muy confiable, pero hacer este análisis en un enunciado con muchos segmentos sería excesivamente lento. Para resolver este problema, los programas se valen de métodos automatizados para obtener de manera rápida y eficiente los valores buscados. El método más usado es el de autocorrelación que, básicamente, consiste en el análisis de una porción fija de tiempo (ventana de análisis), en el que se almacenan los valores de amplitud (los valores en el eje vertical) del oscilograma en esa ventana de análisis. Esa ventana se desplaza a través de la señal y cuando se produce una correlación alta entre los valores almacenados de una ventana anterior y la actual, el programa detecta periodicidad y crea así una curva de una manera rápida y bastante confiable, aún cuando puede tener errores, dado que realiza la estimación a partir de los valores de correlación entre partes distintas de la señal.

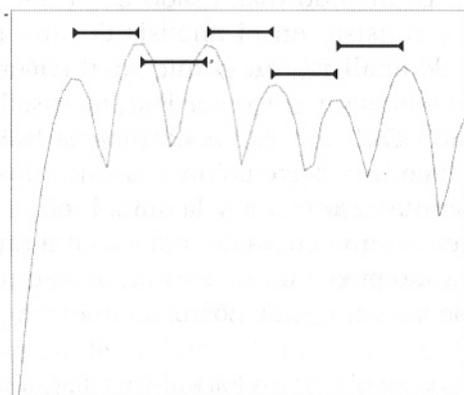
En la figura 3 se observan los valores de f_0 a través del tiempo obtenidos por el método de autocorrelación. Estos valores están en Hz (eje vertical).

Figura 3
Ejemplo de una curva de f_0 obtenida con el método de autocorrelación



Otra manera de estudiar la obtención de un valor de f_0 es mediante el análisis espectral de banda estrecha y observar la relación matemática entre los armónicos. En la figura 4 se presenta una porción de las frecuencias de un espectro de banda estrecha. Las marcas horizontales indican la distancia entre los puntos centrales de cada armónico.

Figura 4
Espectro de banda estrecha que muestra componentes armónicos y su relación frecuencial



Al tratarse de señales periódicas, las relaciones entre los armónicos son constantes y ese valor constante se relaciona directamente con el de la frecuencia fundamental.

En otras palabras, la frecuencia fundamental se puede obtener, por una parte, de manera manual o automática; por otra parte, se puede obtener a partir de un análisis temporal de la señal, o bien, a partir del análisis espectral o frecuencial. Los resultados de ambos métodos dan diferencias irrelevantes.

El método de autocorrelación es el método estándar utilizado actualmente para obtener los valores de la frecuencia fundamental; los otros métodos (medición de la duración del período y análisis espectral) se usan para corroborar el valor obtenido por el método de autocorrelación, en los casos en que se detecte algún error o exista sospecha del valor presentado.

Ahora bien, la curva f_0 es un correlato acústico de las variaciones tonales que se perciben en el enunciado. En la medida en que, por ejemplo, los valores en Hz son más altos, se observa cómo la curva sube y esto se produce porque los pliegues vocales vibran cada vez más rápido. Con respecto al tono se percibiría mayor altura del sonido.

Unidades perceptivas y estilización

Hasta aquí, se ha presentado una definición de tono y se ha señalado cómo se relaciona estrechamente con la frecuencia fundamental y con la vibración de los pliegues vocales. Son distintas caras de una misma moneda.

Un problema interesante en el estudio del tono es que las diferencias en Hz no se relacionan de manera lineal con las diferencias perceptivas. Así, entre 100 Hz y 200 Hz se percibe una diferencia muy grande, de una octava musical. Esa diferencia se produce con 100 Hz de diferencia. Sin embargo, entre 1,000 Hz y 1,100 Hz hay una diferencia perceptiva muchísimo menor a pesar de que hay también 100 Hz de diferencia. Lo que ocurre es que las diferencias en Hz tienen distintos efectos perceptivos dependiendo en qué rango de frecuencias se producen. Mientras más altas son las frecuencias, más Hz se requieren para producir la misma diferencia perceptual que en las frecuencias bajas, en las que se produce con una cantidad menor de Hz. Básicamente, cuando se duplican los Hz percibimos una octava (12 st). La misma diferencia

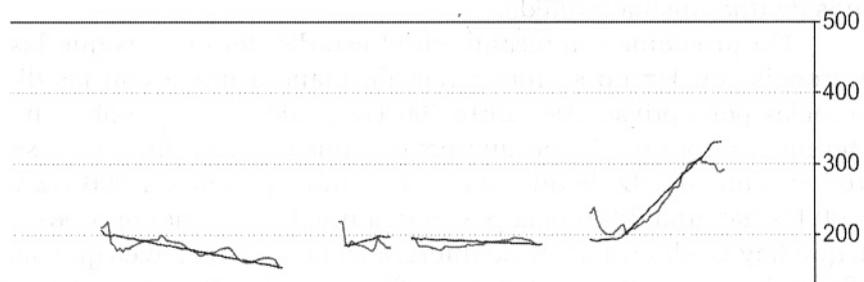
perceptiva se produce entre 100 Hz y 200 Hz, entre 200 Hz y 400 Hz, entre 5,000 Hz y 10,000 Hz o entre 440 Hz y 880 Hz. Cada vez que los Hz se duplican, tenemos 12 semitonos.

Las investigaciones muestran que para el lenguaje usamos un umbral de percepción de diferencias entre 1,5 st (Pamies *et al.*, 2001) y 2 st (t Hart, Collier y Cohen, 1990). En cambio, para la percepción musical, en occidente, la diferencia mínima es de 1 st. En el lenguaje, diferencias entre dos frecuencias que no superen ese umbral no se perciben como distintas tonalmente (aunque sí habrá diferencia en Hz). En otras palabras, habrá diferencias acústicas que no tienen correlato perceptivo.

Estos datos han llevado a la posibilidad de estilizar, o sea, simplificar las curvas de valores de frecuencia fundamental en función de si las diferencias superan o no el umbral perceptivo. La prueba para saber si una estilización está o no bien realizada es la perceptiva: si el sonido con la curva "real" y el sonido con la curva simplificada son indistinguibles, entonces la estilización es funcional a la investigación.

Figura 5

Dos curvas de frecuencia fundamental. Una obtenida por el método de autocorrelación y la otra, una estilización de aquella. Valores verticales en Hz



En la figura 5 se muestran superpuestas dos curvas de tono de una misma frase ("¿Le gusta el pan amasado?"). La que consta de líneas rectas es la curva estilizada y, la otra, es la curva obtenida por el método de autocorrelación. Al escuchar los audios correspondientes a ambas, no se percibe diferencia alguna.

La estilización es una herramienta teórica y técnica muy potente pues permite observar movimientos gruesos del tono y justifica evitar análisis de variaciones microprosódicas irrelevantes para la percepción.

Así, las diferencias en Hz se pueden expresar en términos de st. Hay que tener presente que siempre la medición en st es una medida diferencial y relacional. Ésta sólo se puede establecer al comparar dos frecuencias o un grupo de valores frecuenciales con un valor de referencia. La fórmula para el cálculo de st es:

$$st = ((12/\text{LOG}_{10}(2)) * (\text{LOG}_{10}(\text{frecuencia1}/\text{frecuencia2})))$$

en la que frecuencia1 y frecuencia2 son los valores que se comparan (Martínez Celdrán y Fernández Planas, 2013).

Entre hombres y mujeres hay diferencias naturales de altura tonal determinado por el grosor y la longitud de los pliegues vocales. Los niños presentan valores diferentes a los adultos, tienen un tono más agudo o alto. Evidentemente, las mediciones en Hz son insuficientes para una descripción rigurosa a la hora de comparar hablantes de diferente tésitura, pues ya se ha explicado que las diferencias acústicas no tienen una relación lineal con la percepción. Por esta razón es preferible expresar las diferencias de frecuencias entre distintos hablantes en st.

Hasta aquí hay dos principios fundamentales que darán pie a una propuesta metodológica: la conveniencia de estilizar la curva de frecuencia fundamental y la necesidad de expresar las diferencias frecuenciales en valores perceptivos de st.

Distintas aproximaciones al estudio del tono

Antes de avanzar a la propuesta, es necesario tener presente varias investigaciones que fundamentan lo que aquí se sintetiza o propone.

Así, están los estudios de tpo (t Hart *et al.*, 1990) en los que investigaron meticulosamente la percepción lingüística de variaciones frecuenciales; una buena síntesis se encuentra en Nootboom (1999). Asimismo, investigaciones específicas para el español, las hicieron Pamies *et al.* (2001). Por su parte, la propues-

ta de Cantero (2002) trabaja con porcentajes, no con st, y de esta manera, estiliza la curva de frecuencia fundamental.

Diversos proyectos de análisis entonacional trabajan con criterios inspirados en estos principios. Uno de los más importantes para la lengua española es el proyecto AMPER (Atlas Multimedia de la Prosodia del Espacio Románico) que pretende realizar un atlas de prosodia de las lenguas románicas (Fernández-Planas, 2005).

En este último proyecto se trabaja con tres repeticiones de cada frase. A su vez, cada frase está construida por palabras agudas, graves y esdrújulas en ciertas posiciones. Se trata, como se ve, de habla totalmente estructurada. Para su análisis, se mide la f_0 en tres puntos por cada vocal silábica en una rutina en MatLab; se hace la síntesis del tono y finalmente, se opera con los promedios de las tres repeticiones (para información más detallada, véase Fernández-Planas, 2005).

La propuesta

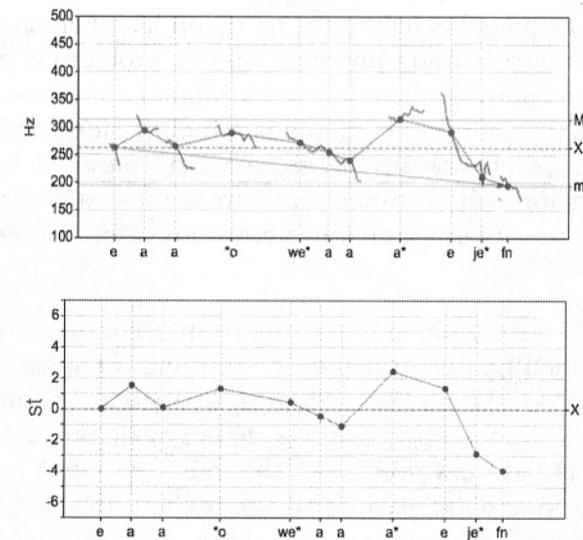
Lo que aquí se propone es un método utilizado en el proyecto chileno "Mapa prosódico de Chile. Descripción fonética y fonológica de rasgos suprasegmentales del español hablado por chilenos", que resuelve, de alguna manera, los problemas antes expuestos. Para la obtención de valores de f_0 se trabaja sólo con las vocales y, además, con el último rastro de la frecuencia fundamental.

Asimismo, se obtiene el promedio de la frecuencia fundamental para cada enunciado analizado y se calcula la diferencia en st de cada uno de los valores en relación con ese promedio. Con ello, un punto que tiene el mismo valor que el promedio, se sitúa en el 0; aquellos puntos que tienen un valor en Hz que da una diferencia positiva en st cuando se comparan con el promedio, se situarán sobre el 0 y a la inversa, valores negativos, bajo el 0.

En la figura 6 se muestra una curva de frecuencia fundamental, su estilización y su expresión en st referidos al promedio.

Figura 6

a) Curva de frecuencia fundamental y estilización a partir de las vocales etiquetadas (arriba); b) desviación en st de cada punto marcado con referencia al promedio (abajo).



Es conveniente hacer una resíntesis de la emisión y sustituir los valores originales por los valores obtenidos en las vocales. La síntesis debe ser indistinguible de la original. Para hacer esto se debe aplicar la opción *Manipulate* -> *To manipulation...* al objeto *sound*. Al editar el objeto *Manipulation* se pueden intervenir los valores de f_0 manualmente. Una manera efectiva es seleccionar toda la señal, eliminar todos los puntos de f_0 (menú superior *Pitch* > *Remove pitch point(s)*) y situarlos en los momentos y en los valores obtenidos (*Add pitch point* e indicar el tiempo y el valor frecuencial). La síntesis nueva se puede escuchar cliqueando en la zona *Visible part* o *Total duration* según lo que se quiera escuchar.

Además se toman en cuenta los valores máximos y mínimos en cada emisión y esa diferencia se expresa en st. También se considera la pendiente que se produce entre el primer punto de f_0 y el último rastro de la frecuencia fundamental para ver una ten-

dencia o línea más general de principio a fin. En la figura 6 a se observan los máximos, mínimo y la recta entre principio y final.

Los materiales de estudio pueden ser emisiones más cercanas a lo espontáneo o más cercanas al habla de laboratorio. El habla más natural presenta problemas a la hora de hacer comparaciones específicas pues los hablantes no dicen las mismas oraciones. El habla más estructurada presenta menos problemas para hacer comparaciones pero se aleja de lo espontáneo.

Lo ideal es combinar métodos e incluir siempre habla más estructurada, como lectura de frases, DCT (*discourse completion task*), descripción o denominación de imágenes, etc. con el propósito de poder hacer comparaciones entre los informantes.

Grabación

Para el proceso de grabación es importante contar con un buen equipo, que facilite, posteriormente, un mejor análisis y la obtención de resultados fiables. En este sentido, se hará alusión a algunas cuestiones técnicas del micrófono y de la grabación como tal.

Respecto al micrófono, se debe evitar el uso del micrófono interno de los computadores. Se recomienda utilizar un micrófono externo, de tipo condensador ya que tiene una mejor respuesta frecuencial, lo que se traduce en una grabación más fiel. Idealmente, debe tener una respuesta frecuencial plana, es decir, que responda con la misma intensidad a las frecuencias entre los 50 Hz y los 10,000 Hz.

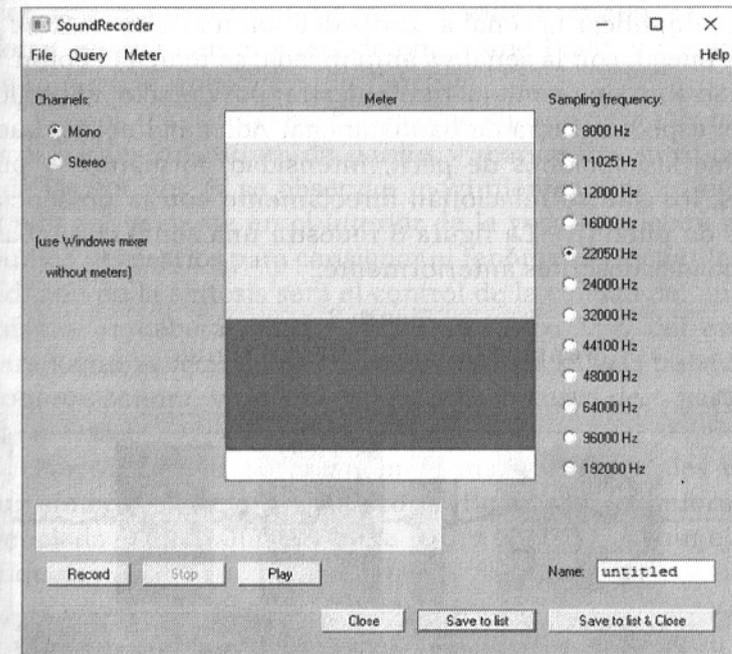
Ahora bien, respecto a la grabación, hay múltiples programas que permiten grabar señales, tales como Praat, Audacity, Wavesurfer, entre otros. En este capítulo, como se ya se ha mencionado, se ha utilizado Praat para ejemplificar los pasos metodológicos propuestos.

Las grabaciones deben ser realizadas en un solo canal, *mono*, salvo que se registre a dos personas por micrófonos distintos. La frecuencia de muestreo (*sampling rate*, *sampling frequency*) puede ser de 22,050 Hz. Dicha frecuencia de muestreo consigna 22,050 muestras por cada segundo en el archivo digital y permitirá analizar frecuencias de hasta 11,025 Hz, lo que es suficiente, dado que para estudiar el habla humana se necesitan frecuencias hasta

los 10,000 Hz y, para la frecuencia fundamental, puede ser incluso menor. Estas opciones se seleccionan en el programa, tal como se presenta en la figura 7, y luego se presiona *Record*. Finalmente, no se debe olvidar guardar el objeto *sound* como archivo *.wav*.

Figura 7

Panel de grabación en Praat (versión para Windows)



Por otra parte, la f_0 es una señal relativamente resistente al ruido y, cierta cantidad de ruido ambiente no impide su obtención con el método de autocorrelación³. Sin embargo, se recomienda que la grabación se haga en un lugar silencioso en el que el ruido ambiente no interfiera con la señal. Además, se debe cuidar que el canal no se sature (*clipping*) para que así, la señal no sea recortada en los puntos de mayor amplitud.

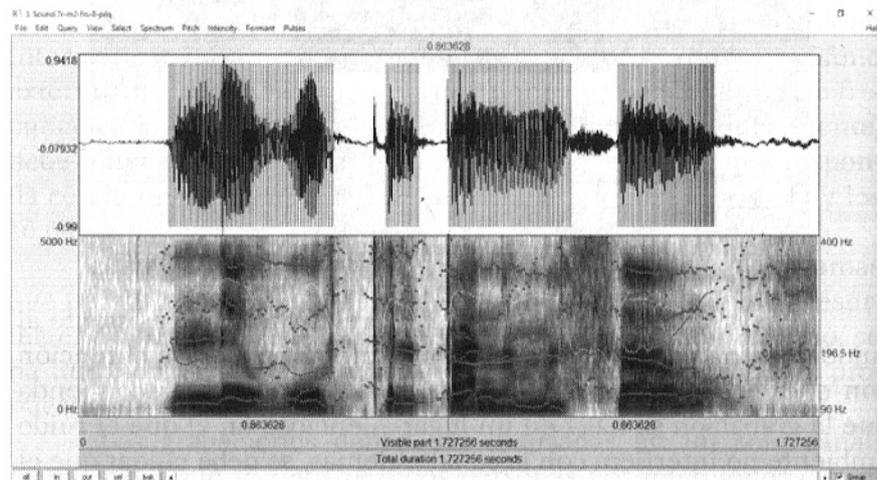
³ Recuérdese que los casos de valores dudosos se pueden cotejar con los otros métodos complementarios.

Edición

En los archivos de audio se deben seleccionar las partes que serán objeto de interés del investigador. Para ello, deben extraerse, uno a uno, los enunciados necesarios creando un nuevo objeto en Praat. Luego, se deben renombrar los archivos con un código que sea informativo para el investigador.

Ahora, con el objeto *sound* seleccionado en Praat, se recomienda amplificar la señal a través del botón *Modify > Scale peak > 0,99*. Luego, con la señal ya amplificada, se marca el botón *View & edit*. Se abre una ventana que muestra, por defecto, el oscilograma y el espectrograma de banda ancha. Adicionalmente, pueden agregarse las opciones de *pitch*, intensidad, formantes y pulsos glóticos, los que se relacionan directamente con la presencia de valores de pitch (f_0). La figura 8 muestra una señal con todas las posibilidades descritas anteriormente.

Figura 8
Ventana de edición de un sonido



Otras de las funciones útiles para la edición de audios en Praat se encuentran en el menú *Edit*. Así, está *cut* para borrar un segmento seleccionado, *paste* para pegar un fragmento de la señal seleccionado anteriormente y *set selection to zero* para convertir en silencio un segmento seleccionado.

Con estas herramientas de edición, el investigador debe limpiar el material que será analizado y quedarse con aquellos enunciados que va a estudiar.

Etiquetaje

Después de grabada y editada la señal, ésta se debe etiquetar con el fin de guardar el análisis elaborado. Praat permite dos opciones: etiquetas de punto o etiquetas de intervalo. Para ello, se debe seleccionar el audio en el panel de objetos de Praat, y elegir, en el menú dinámico, *Annotate > To TextGrid*.

Para el análisis del *pitch*, tal como aquí se propone, hay que crear al menos un estrato de puntos y marcar el centro de cada una de las vocales. Si se observan movimientos tonales significativos perceptivamente en el interior de la vocal, se deben marcar los puntos necesarios para consignar el fenómeno (recuérdese que la audición de la síntesis será el control de la calidad del análisis). Asimismo, se deberá marcar el último rastro de f_0 del enunciado para lograr acoger valores presentes en las sílabas trabadas por consonante sonora, y poder, así, analizar la curva de f_0 hasta el final de la frase.

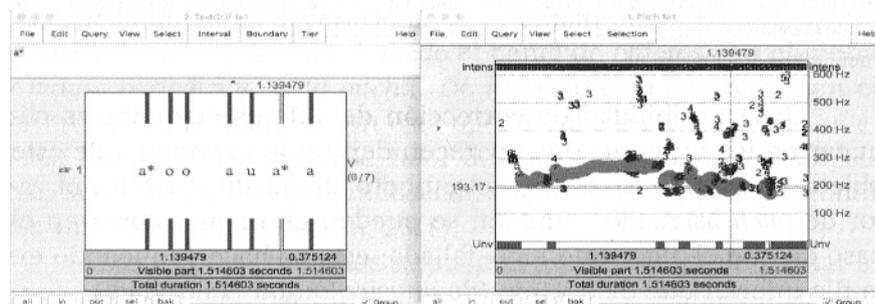
Por otra parte, se recomienda marcar las vocales tónicas con un símbolo, por ejemplo "*", con el fin de obtener información diferenciada para el análisis. En la figura 9 se observa un ejemplo de etiquetaje.

Un ejercicio necesario consiste en escuchar esta síntesis e identificarla como la melodía del enunciado; esto es parte de la evaluación para establecer que el objeto *pitch* está bien editado.

Los valores de frecuencia fundamental se pueden obtener de manera manual o automatizada con un *script*. Hay dos maneras para hacerlo de modo manual.

La primera opción es abrir la ventana de edición para el *TextGrid* y también para el objeto *pitch*. Teniendo sólo estos dos objetos abiertos (esto es importante) se deben agrupar las ventanas. Para ello, se debe activar en ambos objetos, el recuadro *Group* que está en la esquina inferior derecha de la ventana de edición, tal como se puede apreciar en la figura 11. Cuando esta opción está activa, toda selección con el cursor que se haga en uno de los objetos se mostrará también en el otro. Así, habrá que posicionar el cursor en cada marca del *TextGrid* y paralelamente se tendrán los valores de *pitch* en el objeto correspondiente, tal como se observa en la figura 11.

Figura 11
Ventanas de *TextGrid* y objeto *pitch*



Una segunda manera de obtener estos valores manualmente es identificar el tiempo donde está cada etiqueta en el *TextGrid*. Para ello, se puede editar el objeto, situar el cursor en cada marca y registrar el tiempo y la etiqueta; o bien, se puede seleccionar el objeto en el panel de objetos y usar el comando *Query- > Query point tier > Get time of point...* Se señala el número del punto (1, 2, 3, etc.) y se obtiene el tiempo exacto de esa etiqueta. Con el comando *Query- > Query point tier > Get label of point...* se obtendrá la etiqueta en ese tiempo.

Al aplicar estos dos comandos hay que señalar el número de *tier* (1 en este caso) y número de punto (de 1 hasta el número correspondiente a la última etiqueta puesta). El paso siguiente es seleccionar el objeto *pitch* en el panel de objetos y usar el comando *Query- > Get value at time...* y completar ese formulario con el tiempo que se obtuvo en el paso anterior. El resultado será el valor en Hz de la f_0 en ese punto preciso. En esta opción se deben conservar los otros valores por defecto. Este último modo de obtención de los valores se puede automatizar y esto permite obtener rápidamente gran cantidad de datos. Para ello, se debe tener el objeto *Pitch* bien editado y el *TextGrid* correspondiente, con las etiquetas en un estrato de puntos.

El *script* siguiente hace automáticamente la misma tarea realizada manualmente antes.

```
# Entrega valores de  $f_0$  correspondientes a cada etiqueta
# presente en un tier de puntos.
#
# Adaptado de uno original de J.J. Atria para el Laboratorio
# de Fonética de la Universidad de Santiago de Chile
#
# Requiere una selección con dos objetos: un TextGrid,
# y un Pitch que le corresponda
#
# Este script es software libre: puede ser distribuido
# y/o modificado bajo los términos de la licencia
# GNU General Public License publicada por la Free Software
# Foundation, ya sea en su versión 3 o (según se prefiera)
# cualquier versión posterior.
#
# Este script se distribuye con la intención de que sea útil,
# pero SIN GARANTÍA ALGUNA; ni siquiera con la garantía
# implícita de COMERCIABILIDAD o de PROPIEDAD PARA UN FIN
# PARTICULAR. Más detalles pueden encontrarse en el texto de
# la GNU General Public License.
#
# El script debería haber sido distribuido con una copia de la
# GNU General Public License. Si no, este puede encontrarse en
# < http://www.gnu.org/licenses/ >.
#
# Copyright 2016 Domingo Román, Diana Muñoz, José Atria
```

```
tier = 1
precision = 2
```

```

pitch = numberOfSelected("Pitch")
textgrid = numberOfSelected("TextGrid")

if !pitch or !textgrid
  exitScript: "Debes tener seleccionados un objeto Pitch y un TextGrid",
  ... newline$
endif

pitch = selected("Pitch")
pitch$ = selected$("Pitch")
textgrid = selected("TextGrid")

# El primer estrato del TextGrid debe ser de punto
selectObject: textgrid
es_de_intervalo = Is interval tier: estrato
if es_de_intervalo
  exitScript: "El estrato ", estrato, " debe ser de punto",
  ... newline$
endif

# Escribe el encabezado de la tabla a la ventana Info
writeInfoLine: "vocal",
... tab$, "tiempo (s)",
... tab$, "Hz"

# Calcula e imprime resultados para cada punto etiquetado
puntos = Get number of points: estrato
for i to puntos
  selectObject: textgrid
  tiempo = Get time of point: estrato, i
  etiqueta$ = Get label of point: estrato, i

  selectObject: pitch
  f0 = Get value at time: tiempo, "Hertz", "Linear"

  appendInfoLine: etiqueta$,
  ... tab$, fixed$(tiempo, precision),
  ... tab$, fixed$(f0, precision)
endifor

selectObject: textgrid, pitch

```

Alcances y límites de la propuesta

Esta propuesta pone énfasis en la combinación del aspecto perceptivo y acústico, siguiendo la idea común, extendida en la mayoría de los enfoques contemporáneos, de una necesidad de complementariedad entre los diversos puntos de vista del lenguaje.

En el análisis mismo, se propone identificar un valor por cada vocal silábica, lo que se considera suficiente para la mayoría de los casos. Las síntesis de las emisiones confirman que con este método se puede hacer una copia estilizada que cumpla con el requisito de la identidad perceptual con el original.

Los datos que se pueden elaborar con esta metodología son variados y dependerán de los objetivos de cada investigación. En el proyecto para el cual se creó esta propuesta se usan, entre otros, el promedio de f_0 del enunciado, los valores máximos y mínimos de f_0 , la diferencia en st entre el valor máximo y el mínimo, la diferencia en semitonos entre cada punto analizado y el promedio, la diferencia en st entre el primer punto del análisis y el último rastro de la frecuencia fundamental.

Con todo, el análisis que se realiza es de carácter fonético y no aspira a explicar la fonología de la entonación; es decir, es una propuesta para analizar la melodía del habla, más que las propiedades fonológicas de los enunciados desde el punto de vista de la prosodia.

Bibliografía

- Andrews, G.; Ash, R.; Busleiman, A.; Crook, J.; Dannenberg, R.; Daulton, S. y Wharrie, B. (2016). Audacity (Versión 2.1.2) [Software] Recuperado de <http://www.audacityteam.org/>
- Boersma, P. y Weenink, D. (2016). Praat (Versión 6.0.19) [Software] Recuperado de www.praat.org
- Cantero, F. J. (2002). *Teoría y análisis de la entonación*. Barcelona: Edicions Universitat Barcelona.
- de Cheveigné, A. (2006). Multiple f_0 estimation. En Wang, D.L, Brown, G.J. (Eds) *Computational Auditory Scene Analysis, Algorithms and Applications* (pp. 45-79). IEEE Press/Wiley,
- Fernández-Planas, A. (2005). Aspectos generales acerca del proyecto internacional «AMPER» en España. *Estudios de Fonética Experimental*, 14, pp. 13-27.
- Johnson, K. (2011). *Acoustic and Auditory Phonetics* (3ra ed.). Oxford: Wiley-Blackwell.
- Ladefoged (2003). *Phonetic data analysis: An introduction to instrumental phonetic fieldwork*. Oxford: Blackwells.
- Martínez Celdrán, E. y Fernández Planas, A. M. (2013). *Manual de fonética española*. Barcelona: Ariel.
- Pamies, A.; Fernández Planas, A. M.; Martínez Celdrán, E.; Ortega, A. y Amorós, M. C. (2001). Umbrales tonales en español peninsular. En J. Díaz (Ed.), *Actas del II Congreso de Fonética Experimental* (pp. 272-278). Sevilla: Universidad de Sevilla.
- Román, D. (2011). *Manual para el análisis fonético acústico*. Santiago de Chile: Pfeiffer.
- Sjölander, K. y Beskow, J. (2016). WaveSurfer (Versión 1.8.8) [Software] Recuperado de <http://www.speech.kth.se/wavesurfer/>
- 't Hart, J.; Collier, R. y Cohen, A. (1990). *A perceptual study of intonation. An experimental-phonetic approach to speech melody*. Cambridge University Press, Cambridge.