

PROTOCOLO PARA LA EXTRACCIÓN DE DATOS TONALES Y CURVA ESTÁNDAR EN ANÁLISIS MELÓDICO DEL HABLA (AMH)

Miguel Mateo Ruiz
Laboratori de Fonètica Aplicada
Universitat de Barcelona

Resumen

Este trabajo se centra en el desarrollo de una herramienta que facilite a los investigadores el estudio de fenómenos relacionados con la frecuencia fundamental de los sonidos (F0). Concebido en el marco del modelo *Análisis Melódico del Habla (AMH)*, puede no obstante utilizarse con otro tipo de notaciones prosódicas. Presentamos, pues, el programa desarrollado en PRAAT, los resultados de su validación experimental y el protocolo para su utilización. La automatización de la extracción de datos ha de permitir centrar los esfuerzos en el análisis, descripción e interpretación del fenómeno entonativo, facilitando de esta forma su enseñanza y aprendizaje.

PALABRAS CLAVE: entonación, análisis melódico, extracción automática de datos tonales

Resum

Aquest treball presenta una eina que permeti als investigadors l'estudi de fenòmens fònics relacionats amb el paràmetre de la freqüència fonamental dels sons (F0). S'ha fet en el marc del model *Anàlisi Melòdic de la Parla (AMP)*, però aquest mètode pot ser utilitzat amb altre tipus de notacions prosòdiques. Presentem, doncs, el programa que hem desenvolupat en PRAAT, els resultats de la seva validació experimental i el protocol per poder utilitzar-lo. Aquesta automatització de l'extracció de dades tonals ha de permetre que els esforços dels investigadors es centrin en l'anàlisi, descripció e interpretació del fenomen entonatiu, facilitant doncs tant l'ensenyament com l'aprenentatge.

PARAULES CLAU: entonació, anàlisi melòdic, extracció automàtica de dades tonals

Abstract

This paper presents a tool that allows researchers to study phenomena related phonic parameter of the fundamental frequency of sounds (F0). Has been made under the model "*Melodic Analysis of Speech (MAS)*", and can be used with other prosodic notations. We present then the program developed in PRAAT, the results of its experimental validation and the protocol to use the program accurately. The automation of data extraction must allow focusing on the analysis, description and interpretation of intonational phenomenon, thereby facilitating teaching and learning.

KEYWORDS: intonation, melodic analysis, tonal data automatic extraction

1. Introducción

El estudio de la entonación ha experimentado un notable crecimiento desde las últimas décadas del siglo XX y en la actualidad existen diversos grupos de investigación que trabajan en la confección de atlas prosódicos, en la comparación de la prosodia de catalán y castellano. Síntoma de este creciente interés son también los diferentes congresos y seminarios monográficos que han tenido lugar en los últimos años.

Esto ha sido posible por el desarrollo de nuevos medios instrumentales y software de obtención de F0 cada vez más rápidos y eficaces.

El objetivo de nuestra investigación es el desarrollo de una herramienta fiable y sencilla de utilizar que permita automatizar el proceso de recogida de datos para el estudio de la entonación de cualquier lengua, aunque inicialmente se ha diseñado y construido con los parámetros del *Análisis Melódico del Habla (AMH)* que son válidos para el castellano y el catalán.

1.1. El modelo AMH

El modelo AMH, expuesto de forma detallada en Cantero (2002) y Font (2007) distingue tres tipos de funciones o niveles de la entonación:

- Entonación *lingüística*: de identidad fonológica, significativa, definida mediante la oposición de tres rasgos: interrogativo, enfático y suspendido. La entonación lingüística es la interpretación fonológica de la melodía.
- Entonación *prelingüística*: la entonación es un elemento clave tanto en la integración como en la fragmentación y comprensión del discurso – Quilis (1981), hablaba de la “función integradora, delimitadora” de la entonación- .
- Entonación *paralingüística*: expresiva, sirve para indicar actitud, el estado de ánimo del interlocutor, etc.

La entonación se articula mediante procesos dinámicos de acuerdo a una determinada jerarquía fónica (entonación prelingüística) para producir entidades significativas, fonológicas (entonación lingüística). Estas representaciones no son cerradas, permiten la incorporación de muy diversos sentidos y significaciones no lingüísticas (entonación paralingüística).

El modelo se basa en el concepto de *jerarquía fónica*: cuando hablamos, lo hacemos de forma jerarquizada, no enlazamos sonidos de forma continua (el tradicional concepto de “cadena fónica”) sino que a través de los fenómenos suprasegmentales agrupamos los sonidos en diferentes bloques.

Esta jerarquía se da en varios niveles: sílabas, palabras fónicas y grupos fónicos. Los tres constituyen bloques de sonido que se agrupan en torno a un acento. Los únicos elementos tonalmente significativos son las vocales, los segmentos que aparecen acústicamente marcados y destacados en el análisis de la línea melódica.

Los niveles de jerarquización son los siguientes:

- Sílaba: la vocal es su núcleo, siguiendo la distinción tradicional entre vocales y consonantes.
- Palabra fónica: algunas vocales son más relevantes, cuando hablamos hay vocales que son más perceptibles que otras, son las vocales tónicas. Estas vocales son siempre el núcleo de una palabra fónica y en torno a ella pueden agruparse otros elementos átonos, que pronunciamos de forma conjunta; generalmente son elementos gramaticales.
- Grupo fónico: no hablamos con sucesiones de palabras. En nuestro discurso, organizamos una serie de palabras (**fónicas**), en torno a un núcleo, el acento sintagmático que se materializa en una inflexión tonal. El grupo fónico coincide con lo que denominamos, desde el punto de vista de la entonación, como contorno entonativo: la sucesión de tonos de las vocales organizados en torno a una inflexión final.

1.2. El método de análisis

El método de análisis *formal* que se propone está basado en el análisis acústico y perceptivo del discurso. Esto permite hacer una descripción de la entonación desde un punto de vista fonético. Veamos, paso a paso, el procedimiento del método de *Análisis Melódico del Habla*:

El primer paso es la identificación de las unidades melódicas de los enunciados de habla espontánea (de los grupos fónicos); a diferencia de otros estudios sobre la entonación realizados con habla de laboratorio, en contextos predeterminados por el investigador, que con frecuencia además es también el único informante del experimento, con lo cual no se estudia la realidad sino la realidad inducida por los planteamientos del investigador.

El método de Cantero permite el análisis de las unidades fónicas del habla independientemente de otros niveles de análisis (gramaticales, pragmáticos, etc.), permite trabajar con habla espontánea –como reclamaba ya Navarro Tomás hace más de 70 años-, con informantes anónimos que ignoran que lo son.

De acuerdo a los presupuestos teóricos de la jerarquía fónica, realizaremos la segmentación melódica con un criterio formal, según las inflexiones tonales, la melodía de cada grupo fónico vendrá determinada por la F0 de los segmentos tonales (las vocales) que los componen.

El segundo paso es la fase de recogida de datos acústicos relevantes de los diversos segmentos tonales y su estandarización: esta recogida de datos se hace, en nuestro caso, mediante PRAAT¹, un software de análisis acústico fiable. Para ello, se identifican las vocales y los segmentos tonales que en ella se producen, apoyándonos en el sonograma, y se recoge el valor medio –si es estable- o el valor central. En el caso de vocales tónicas puede producirse una inflexión y existir dos y hasta tres segmentos tonales –inflexión circunfleja-; estos valores se toman de los extremos de la inflexión. Esta

¹ <http://www.fon.hum.uva.nl/praat/>. Se ha utilizado la versión 5.1.32

A continuación se realiza la estandarización de los valores obtenidos, como dijimos la melodía no es una mera sucesión de sonidos, sino que existe una relación entre ellos, no son por lo tanto significativos sus valores absolutos sino la sucesión de valores relativos: una misma diferencia en hercios entre dos segmentos tonales, 60 por ejemplo, será diferente según el valor del segmento tonal inicial, si es 120, la variación tonal es de un 50%; si el valor inicial es 240, la variación tonal es menor, únicamente de un 25%.

Esta estandarización, se realiza mediante una sencilla regla de tres, se asigna 100 al primer valor de la frase y a partir de él se calcula la curva melódica según las variaciones tonales en porcentaje y obteniendo unos valores estandarizados partiendo del 100 inicial. Otros modelos realizan este proceso en semitonos con la consiguiente complicación del algoritmo de cálculo y pensamos que una pérdida de claridad, nuestro sistema es más intuitivo, permite al investigador un análisis más sencillo al estar expresado de una forma lineal. El resultado de esta estandarización es la expresión de la melodía del enunciado, veámoslo en la figura 3, siguiendo con el ejemplo anterior.

Enunciado	Pe	ro	sa	bes	bai	Lar	lar*
Hercios	227	233	264	354	357	411	476
Porcentaje	100%	2,6%	13,3 %	34,1 %	0,8%	15,1%	15,8%
Curva estándar	100	103	116	156	157	181	210

Figura 3 – Tabla con los datos estandarizados

Los valores estandarizados de la curva permiten representar gráficamente la curva y la posterior clasificación y comparación

AN-26-02-08

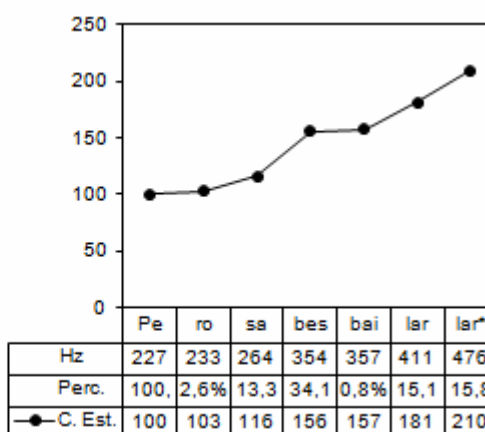


Figura 4 – Curva estandarizada del enunciado AN-26-0208.

La estandarización y el gráfico los realizamos mediante la grabación de los datos en una herramienta que lo permita, Excel, en nuestro caso (v. figura 4).

Como decíamos, al finalizar esta fase ya podemos clasificar los contornos, con la gran ventaja de que lo podemos hacer independientemente de cualquier variable tanto del informante (sexo, edad) como de la duración de los grupos fónicos o de su gramaticalidad ya que al haber normalizado y relativizado los valores (en base al arbitrario 100 inicial) es posible realizar la comparación de valores exclusivamente melódicos. Veámoslo con un ejemplo:

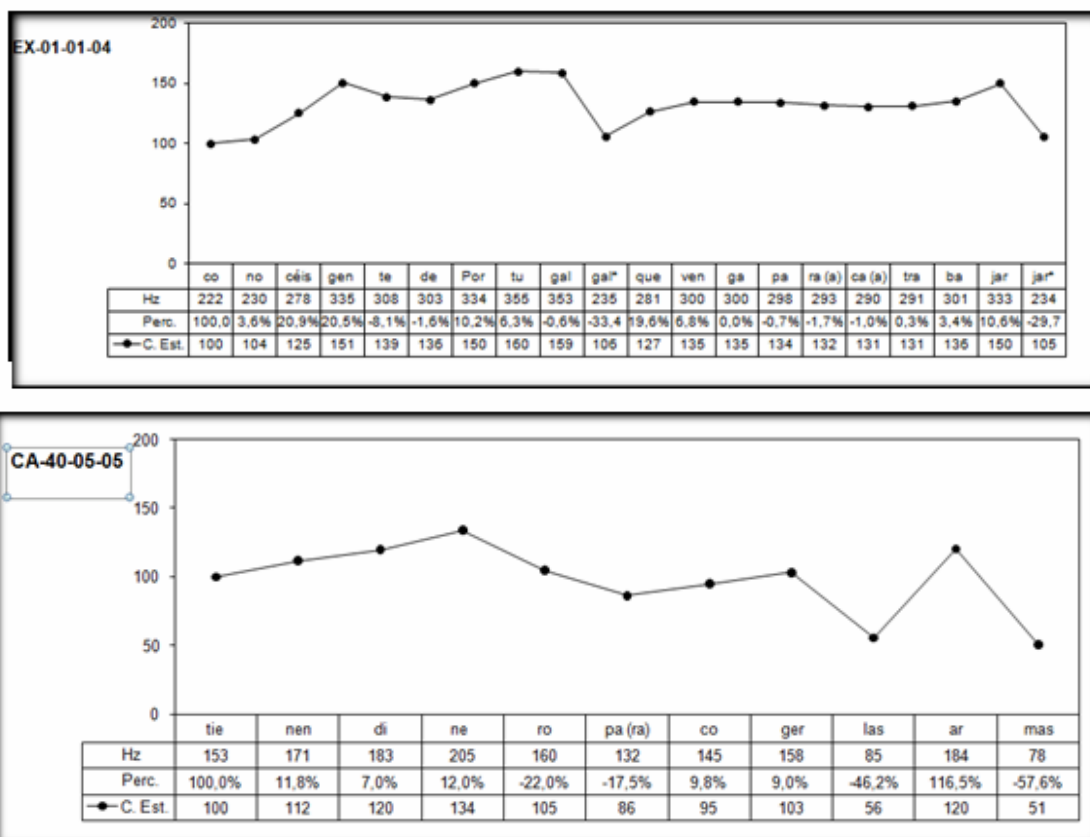


Figura 5 – Curva estándar de los enunciados EX01-01-04 y CA-40-05-05

Los dos enunciados (figura 5) pertenecen al mismo patrón melódico, su melodía, con una inflexión final situada al mismo nivel del primer pico, es la misma desde un punto de vista estrictamente melódico, a pesar de ser de dos informantes diferentes –mujer el primero, hombre el segundo-, con longitudes diferentes y cada una con su significado.

Tras la fase acústica, los resultados se pueden validar mediante una prueba perceptiva. PRAAT permite obtener una copia sintetizada del enunciado original (método PSOLA) y después podemos sustituir los valores tonales por los estandarizados y realizar tanto la validación perceptiva como el establecimiento de sus valores fonológicos, identificando los valores significativo y sus márgenes de dispersión, ya que mediante la manipulación y prueba perceptiva posterior podemos establecer qué cambio, qué valor hace que un rasgo (fonológico) pase de /-enfático/ a /+enfático/ por ejemplo.

Estas variaciones se pueden producir en cualquiera de los elementos funcionales del contorno –los rasgos melódicos- :

- Anacrusis: sílabas átonas anteriores a la primera tónica del contorno.
- Primer pico: primera vocal tónica.
- Cuerpo (declinación): sílabas que van desde el primer pico hasta la última vocal tónica, dónde empieza la
- Inflexión final: sílabas desde la última vocal tónica al final del grupo fónico.

2. Objetivos de la investigación

Como hemos apuntado, nuestra investigación se inscribe en el contexto de la constatación de la importancia del aprendizaje de la entonación en la enseñanza de lenguas y en la necesidad de una descripción previa de los diferentes patrones melódicos de las lenguas, según la teoría del *Análisis Melódico del Habla*. Nuestra contribución pretende ser un primer paso en el camino de poner a disposición de los investigadores una herramienta robusta, fiable que permita superar las resistencias de aquellos que como LL. Manchón (2008:84) o A. Cabedo están convencidos de la bondad del método y de su coherencia teórica pero manifiestan algún pero: “no obstante cabe señalar que se trata de un procedimiento costoso que, de momento, únicamente puede realizarse de modo manual, puesto que es complicado desarrollar un algoritmo que realice las funciones de estilización mencionadas: buscar el contorno de la vocal, encontrar su valor frecuencial, encontrar el porcentaje de ascenso o descenso tonal (sería la parte menos conflictiva de alcanzar una vez conseguido lo anterior), etc.,” (2009:89)

Este objetivo global lo podemos reformular en tres:

- Desarrollar un programa para PRAAT que permita la extracción semiautomática de datos tonales (segundo paso del protocolo descrito en el apartado anterior, la recogida de datos relevantes y estandarización).
- Establecer su fiabilidad: verificar experimentalmente la bondad de los resultados obtenidos.
- Definir el protocolo de uso: establecer y describir, de forma clara y detallada, los pasos a seguir para la nueva forma de obtención de datos tonales en el método de análisis (AMH) que utilizamos.

3. Metodología

La viabilidad de nuestro primer objetivo está determinada por el *software* de análisis acústico que de manera generalizada se utiliza en la investigación fonética. PRAAT es una herramienta que se utiliza en el ámbito científico para estudios de las diversas características acústicas del habla, entre ellos la entonación, gracias a los datos de la frecuencia fundamental (F0) que extrae mediante algoritmos robustos, con resultados validados y aceptados. Entre sus ventajas, además de permitir la manipulación y síntesis de los enunciados para poder realizar pruebas perceptivas –básicas en nuestra metodología-, hay que destacar la que nosotros hemos utilizado en nuestro trabajo: dispone de un lenguaje de programación propio que permite sistematizar y automatizar ciertas actividades de forma que no es necesario hacer la toma de datos manual (medir, apuntar, trasladar a una hoja Excel, etc.), un proceso laborioso. Con estos programas, se puede realizar la extracción de forma automática, generando la información en archivos para su posterior estudio y análisis.

Son numerosos los sitios web en los que hay disponibles programas (*script* en terminología PRAAT) desarrollados por otros investigadores para extraer información sobre duraciones vocálicas, formantes, frecuencia fundamental, etc.; a ellos hemos acudido para hacernos una idea del lenguaje y extraer ideas sobre cómo realizar el nuestro².

Así pues, en nuestro estudio lo que hemos hecho es codificar un programa en PRAAT que extrajera los datos de F0 de los segmentos tonales de un enunciado según los criterios del protocolo de *Análisis Melódico del Habla* descritos en el apartado anterior.

El segundo objetivo es establecer la fiabilidad del programa, para ello hemos hecho la verificación experimental: tras realizar el análisis manual de los datos, procesamos el *script* desarrollado y verificamos los datos.

En una primera etapa hemos comparado los resultados y su grado de coincidencia (análisis cuantitativo). Una vez depurado y mejorado el programa, hemos analizado aquellos casos en que la diferencia entre los datos obtenidos manualmente y los datos extraídos con el programa es superior al 10% (análisis cualitativo).

Para realizar esta verificación experimental hemos utilizado un corpus de habla espontánea de 80 oraciones interrogativas absolutas del castellano –ver anexo 1-. Hemos utilizado este corpus ya que aunque en nuestro estudio no pretendemos describir ningún patrón entonativo concreto, estamos realizando en el LFA una investigación sobre este tipo de enunciados en el que las mediciones de los segmentos tonales se han realizado manualmente y han sido validados por dos investigadores por lo que constituyen una referencia sólida con la que contrastar los datos obtenidos mediante el programa que hemos desarrollado.

² Citamos, a modo de ejemplo, (<http://www.helsinki.fi/~lennes/praat-scripts> y http://ed268.univ-paris3.fr/lpp/pages/EQUIPE/gendrot/page_web/scripts.htm) ya que pensamos que el manual de programación de PRAAT es poco usable sin un conocimiento previo. Se puede localizar un amplio catálogo de direcciones en <http://www.linguistics.ucla.edu/faciliti/facilities/acoustic/praat.html> o en foros de discusión como <http://queenoflily.blogspot.com/2008/11/praat-script-resources.html> .

Es un corpus extraído fundamentalmente del corpus de variedades dialectales del castellano que hemos elaborado en el Laboratorio de Fonética Aplicada³. Los informantes son 72 (45 mujeres y 27 hombres) con edades comprendidas entre 16 y 80 años. Los 80 enunciados analizados contienen un total de 796 segmentos tonales.

4. El script

El objetivo de los programas es realizar de forma semiautomática dos de las fases del método: la captura de datos y la estandarización de la curva melódica. Decimos semiautomática porque nuestra experiencia al analizar manualmente enunciados de habla espontánea, genuina es que debido a las características tanto de las voces como del propio discurso oral, en ocasiones PRAAT no puede extraer todos los datos tonales (F0) y éstos deben ser calculados manualmente (a partir del número de ciclos por período de tiempo) o extrae valores que son el 50 o el 100% de los reales. Por ello es precisa la intervención del investigador.

Antes de describirlos, señalaremos que se han concebido para ser utilizados dentro de un proceso de pasos sucesivos ya que nuestras investigaciones se basan en el estudio de volúmenes significativos de datos de habla espontánea que permitan la validación y generalización de las conclusiones. A modo de ejemplo, el corpus de variedades dialectales al que hemos hecho referencia anteriormente, está constituido por más de 2500 enunciados. Estos pasos, que se describen ampliamente más abajo, en el apartado 6 (protocolo), son los siguientes:

a) *Segmentación y etiquetaje de los segmentos tonales.*

Es el punto clave de todo el proceso, el etiquetaje ha de ser preciso, el investigador ha de realizar este proceso con suma atención y la máxima precisión, recurriendo a la ampliación del espectrograma (botón “SEL” de la pantalla con sonido más *textrid* en que realizamos la segmentación y etiquetaje), modificación de los límites del rango de hercios, (en la pantalla “*pitch settings*”), atención a la onda sonora además del sonograma, de modo que pueda establecer una correcta segmentación del enunciado; en esta primera versión nuestra herramienta únicamente realiza la extracción de los datos de los segmentos tonales establecidos por el investigador.

b) *Extracción de datos.*

Este paso lo realiza el primero de los scripts construidos.

c) *Revisión de alertas y d) corrección de datos.*

A partir de la información del script anterior, se realizan las correcciones oportunas.

³ Comunicaciones (en prensa) presentadas al XXXIX Simposio Internacional de la Sociedad Española de Lingüística, Santiago de Compostela 2010 (ptes . publicación) : Ballesteros, M, M. Mateo & F.J.Cantero, , *Corpus oral para el análisis melódico de las variedades del español* y Cantero, F.J, M. Ballesteros y M. Mateo: *Proyecto de investigación Análisis Melódico del Habla*, en el marco del proyecto de investigación FFI2009-13214-C02-01.

e) *Confección de la curva estándar.*

Una vez corregidos, si procede, los datos de los que se han creado alertas, se obtiene la curva estándar de los enunciados mediante el segundo de los scripts.

f) *Generación de gráficos con la cura estándar y comparación de modelos.*

El último paso del proceso es generar los gráficos con la curva estándar, tal y como se hace en el protocolo actual, de recogida manual de datos, mediante documentos de Excel⁴ que incorporan el cálculo de la curva estándar y la confección del gráfico.

El primer programa necesita la información que se etiqueta en el apartado (a), el *textgrid*, para realizar el paso de extracción de datos de F0 (b). El segundo programa se procesa una vez revisadas las alertas (c) y modificados los datos tonales que lo requieran (d) o se detecten anomalías en la revisión y comparación de las curvas estandarizadas con los modelos (e).

4.1. El script de extracción de datos tonales

Tras realizar validaciones formales: que el sonido y el *textgrid* a procesar sean el mismo fichero y que en este segundo elemento se haya indicado si la voz es masculina o femenina, se realiza la extracción de los datos de pitch del enunciado (objeto pitch, en terminología PRAAT).

Esta extracción se realiza mediante el método que PRAAT recomienda para el tratamiento de la entonación (*autocorrelation*) y con todos los parámetros con sus valores estándar, excepto tres:

- Rango de hercios: diferenciamos según la voz sea masculina (40-350) o femenina (90-500).
- Método ‘*very accurate*’: mediante una ventana gaussiana, con lo cual la longitud efectiva sobre la que efectúa la extracción del tono es el doble que la que utiliza con los valores estándar.
- *Voicing threshold* : PRAAT recomienda disminuir este valor en casos en que no se detectan correctamente los tramos con voz; alude a casos de voces patológicas o en que los períodos a analizar son irregulares como es nuestro caso al tratarse de habla espontánea. Como veremos al comentar los resultados, su modificación ha sido pertinente.

Una vez hemos extraído el tono del enunciado, analizamos los segmentos tonales que el investigador ha marcado en el *textgrid* para identificar las vocales. De cada segmento tonal extraemos cinco datos: el valor medio (mean), el valor mínimo, el valor máximo, el valor inicial y el valor final. Si alguno de estos dos últimos valores no está informado, realizamos un proceso de búsqueda para encontrar el primero o el último informados dentro de la secuencia temporal del segmento tonal.

⁴ Entre los siguientes objetivos de automatización del protocolo está la generación desde PRAAT de los gráficos. Entre tanto se ofrecerá a los investigadores que trabajan en AMH un programa en Excel que automatiza la confección de los gráficos.

El siguiente paso es determinar cuantos valores tonales significativos hay en el segmento; en el caso del castellano y catalán se ha verificado en diversos experimentos perceptivos realizados en el Laboratorio de Fonética Aplicada de la Universidad de Barcelona⁵ que este umbral es del 10%, por ello verificamos la diferencia existente entre los valores mencionados y en función de dicha diferencia grabamos los valores:

- Un valor: principal, media: si las diferencias entre los valores tonales son inferiores al 10%.
- Dos valores, F0 inicial y F0 final: si las diferencias entre ellos son superiores al 10% y no lo son respecto al máximo y mínimo.
- Dos valores, F0 inicial y F0 mínimo: si la diferencia entre F0 mínimo e inicial es superior al 10% pero no la diferencia entre mínimo y final.
- Dos valores, F0 inicial y F0 máximo: si la diferencia entre F0 inicial y máximo es superior al 10%, pero no la diferencia entre los valores máximo y final.
- Tres valores, F0 inicial, mínimo y final : si las diferencias entre los tres valores son superiores al 10%
- Tres valores, F0 inicial, máximo y final: si las diferencias entre los tres valores son superiores al 10%.

Si no se puede extraer alguno de los valores generamos una alerta para la posterior revisión manual por parte del investigador.

En este proceso el programa genera dos archivos:

- Alertas: sólo se generan para aquellos enunciados en los que se produce alguno de los fenómenos que requieren revisión del investigador
El programa genera una línea para cada segmento tonal del enunciado con el segmento tonal afectado y una clave que identifica el tipo de alerta. En estos momentos están definidas las siguientes:
 - Si no se ha podido encontrar valor para un segmento tonal: se graba el valor 0.
 - Si se detectan valores por encima o debajo de los rangos de voz masculina y femenina que se han definido: se graba el valor 999.
 - Si se detectan tres valores en el segmento tonal. Se graba el valor 3.
- F0: datos tonales de cada segmento del enunciado, es un fichero intermedio que sólo es importante y susceptible de ser utilizado para aquellos enunciados de los que se han creado alertas.

Una vez extraídos los datos tonales de todos los segmentos, realizamos el proceso de estandarización según el procedimiento descrito en el apartado 1.2; al primer segmento le asignamos los valores 100% (variación) y 100 como valor de curva estándar y realizamos la regla de tres descrita anteriormente.

⁵ Consultar página web del laboratorio : <http://www.ub.es/lfa/>

4.2. El script de cálculo de la curva estándar

Este programa lee un fichero con los datos de F0 y realiza la estandarización ya descrita, generando los archivos con la curva estándar, contiene los datos de la estandarización que importaremos a Excel para generar el gráfico con la curva melódica del enunciado.

5. Resultados

En este apartado presentamos los resultados obtenidos en el proceso de elaboración del programa de extracción de datos tonales. La primera versión que realizamos, siguiendo algunos ejemplos del manual de PRAAT, obtuvimos valores válidos en un alto porcentaje pero también un número muy elevado de alertas: 71 valores, un 9,66%, sin informar⁶.

La mayoría de estos datos no informados se daba al obtener el valor inicial o final del segmento tonal, como hemos indicado anteriormente son los valores junto al máximo y mínimo que intervienen en el proceso de decisión del valor tonal que hay que extraer. El problema se produce porque el algoritmo de PRAAT que realiza el muestreo para obtener la información tonal no proporciona información “continua”. Incorporamos entonces una modificación al programa para buscar el primer y el último marcadores con información de tono en cada segmento.

Los resultados mejoraron, únicamente 22 alertas (2,99%) por segmentos tonales sin información, pero apareció también un elevado número de alertas por valores que catalogamos como extremos, por encima o debajo del tono estándar del habla: 40 y 550 hercios, en concreto 15 (1,88).

Son porcentajes pequeños, pero mirando el resto de parámetros que establecimos revisar para la validación (cuantitativa) de los resultados, nos llevaron a la conclusión de que eran necesarios más ajustes. Los parámetros que comprobamos son los siguientes:

- Número de valores tonales obtenidos por el programa.
- Porcentaje de coincidencia con los valores obtenidos manualmente, comprobamos los casos de coincidencia exacta (0% de desviación) y cuatro niveles de diferencia: inferior al 2%, inferior al 5%, entre el 5 y el 10% y superiores al 10%
- Número de alertas: por extracción de 3 valores, por valores extremos y por valores sin informar.
- Valores tonales extraídos por el programa que no se han obtenido manualmente, distinguimos por una parte, aquellas diferencias que son debidas a un mayor grado de precisión en el cálculo automático y por tanto el programa obtiene dos valores con una diferencia en torno al 10% y por otra aquellas que son superiores.

⁶ Estos datos son sobre 735 segmentos tonales ya que al iniciar el script aún no disponíamos de los tres enunciados de la variedad aragonesa (AR-XX-XX-XX, en el inventario del corpus, anexo)

- Valores tonales obtenidos manualmente, que el programa no ha calculado.
- Número de enunciados con el 100% de los valores calculados correctamente.

Las siguientes modificaciones que realizamos surgieron de la reflexión con otros investigadores que nos llevaron a revisar nuestra forma de actuar al realizar la extracción manual de datos. Observamos que dado que nuestros corpus están constituidos por habla espontánea y son muy extensos en cuanto a número y variedad, en ocasiones tenemos que cambiar algunos de los parámetros de configuración del algoritmo de extracción de pitch de PRAAT (v. figura 6).

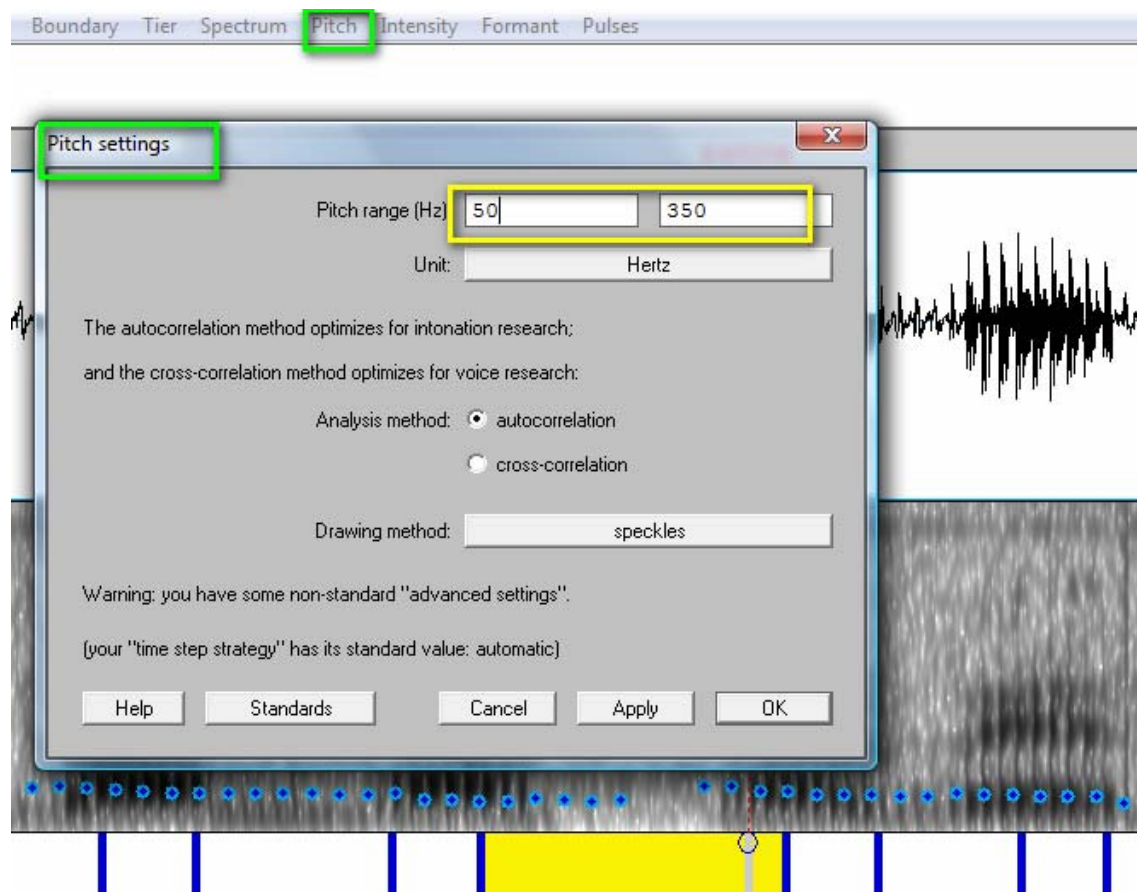


Figura 6 – PRAAT, pantalla de parametrización extracción F0.

El rango estándar está definido entre 75 y 500 hercios pero en ocasiones, por las características de las voces el algoritmo de determinación de la frecuencia fundamental no funciona correctamente sino se adaptan estos rangos, bajándolo para las voces masculinas –más graves- y subiéndolo para las femeninas –más agudas-; así estableciendo rangos de 50 -350 y 90-550 hercios respectivamente se solventan estos problemas.

Decidimos pues, introducir esta variable en nuestro programa; modificamos los *textgrid* para incorporar esta información sobre el informante del enunciado en el primer segmento (*tier*) y los resultados mejoraron globalmente, siendo los enunciados totalmente correctos un 75% del total, 60 enunciados.

En las figuras 7 y 8 observamos que la comparación de los valores tomados manualmente con los extraídos por el programa y sus diferencias, mejoró:

Diferencia	%
0%	29,02%
< 2%	75,50%
<5%	90,20%
5-10%	96,36%
> 10%	3,52%

Figura 7 – Diferencias datos manuales-datos script

Alertas		
3 valores	31	3,9%
Sin informar	0	0%
Valores extremos	6	0,8%

Figura 8 –Alertas

Por el contrario, se produjo un considerable incremento de los valores tonales extraídos por el programa, frente a los 796 que anotamos manualmente para los 80 enunciados, al realizar la extracción automática obtuvimos 939, un 18% más y de ellos, casi un 35% de estos valores no son debidos a la precisión en el cálculo del 10% de variación dentro del segmento tonal.

Decidimos por ello revisar los parámetros de extracción del pitch de PRAAT (“Advanced Settings” en el menú “Pitch”) y tras verificar su uso decidimos cambiar el mecanismo de extracción a “very accurate”, que aumenta la ventana de muestreo, según indicamos en las características del programa. Hicimos entonces dos versiones y verificaciones, una primera con el parámetro “voicing treshold” (VT) con valor estándar y otra reduciendo el valor a 0,25 que recomienda PRAAT para los casos en que no se detectan correctamente los tramos de pitch con voz.

Los resultados (v. figura 9), comparando con los obtenidos manualmente son similares.

	S.Tonales	0%	<2%	< 5 %	5-10%	>10%	Enunciados 100% OK
VT estándar	861	30,28%	80,03%	90,95%	95,35%	4,15%	53
VT 0.25	893	29,02%	77,14%	90,95%	96,61%	2,76%	59

Figura 9 –Comparación resultados con cambio parámetro VT

Podemos apreciar que el grado de coincidencia es prácticamente el mismo, un 95,35% frente a un 96,61%; el dato significativamente peor en el caso de los valores estándar es el porcentaje de casos en que la diferencia entre el valor manual y el valor obtenido mediante el script es superior al 10%; también el número de enunciados en que todos los valores son correctos es un 10% mayor en el caso en que el valor de VT no es estándar. En contraste con la versión anterior, el número de segmentos tonales ha disminuido y ha pasado a ser de un 8% y un 10,2% respectivamente, frente al 18% anterior. Como veremos en la siguiente tabla, la mayoría de los casos son debidos a la precisión al calcular las variaciones del 10%., en torno al 75% en ambos casos.

En cuanto a las alertas (v. figura 10), en las tres primeras columnas tenemos el número de alertas, con resultados similares, excepto en el caso de valores sin informar, con el parámetro VT no estándar (0.25), siempre se extrae valor. Los enunciados en los que con valores estándar no se obtiene el valor son 8 ya que hay un enunciado – MU-98-03-05 – que tiene cuatro valores sin informar. De estos 8 enunciados, los valores extraídos con el parámetro VT no estándar son correctos para tres enunciados (AI-26, AN-85-02-02 y CA-40-05-05), en dos casos hay diferencias por precisión en el cálculo de diferencias tonales del 10% (AI-10 y AN-26-01-06) , otro aún teniendo dos valores en un segmento, la curva melódica no se ve afectada (AN-82-06-01). Los dos enunciados restantes, uno se apreciará el error de la extracción en la verificación visual (valor 100% diferente, enunciado CA-43-03-03) y para el ya referido del enunciado del informante murciano, se ha generado una alerta porque el programa extrae tres valores.

	3 valores	Valores no informados	Valores extremos	Valores 10%	Valores No 10 %	Valores no script
VT estándar	14	11	1	57	18	4
VT 0.25	15	0	3	74	25	5

Figura 10 –Comparación alerta con cambio parámetro VT

Por lo que respecta al mayor número de segmentos tonales respecto a la medición manual y la no captura de datos por el programa (última columna) las diferencias no son significativas; en el caso de valores debidos a una diferencia por precisión en el cálculo al efectuar la captura mediante el programa (columnas 4 y 5), es lógico que sea superior cuando los valores no son estándar porque esta opción hace que se registren más datos de pitch. Veamos el primer caso –columna 4-.

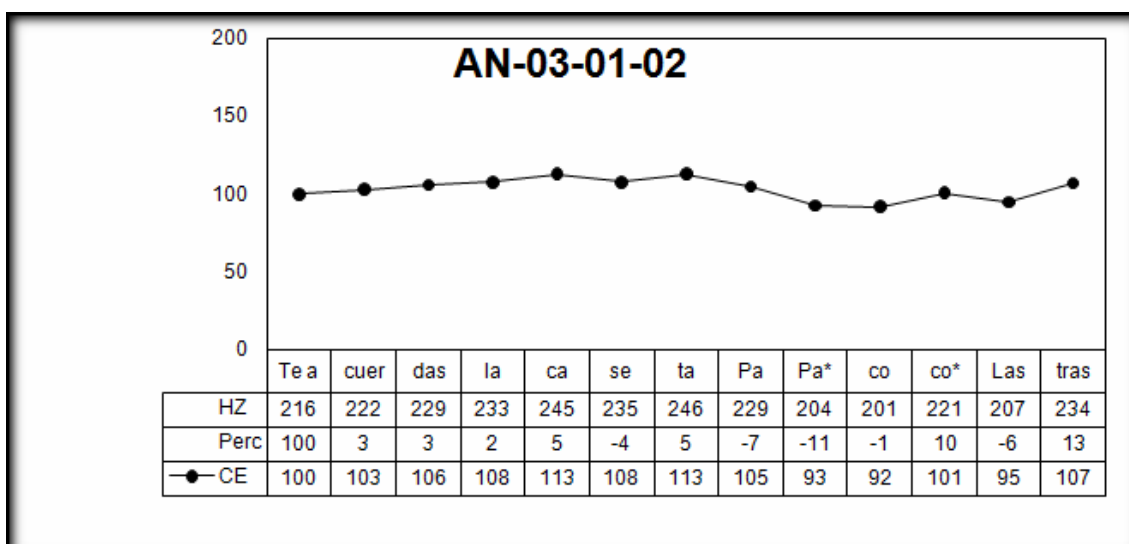
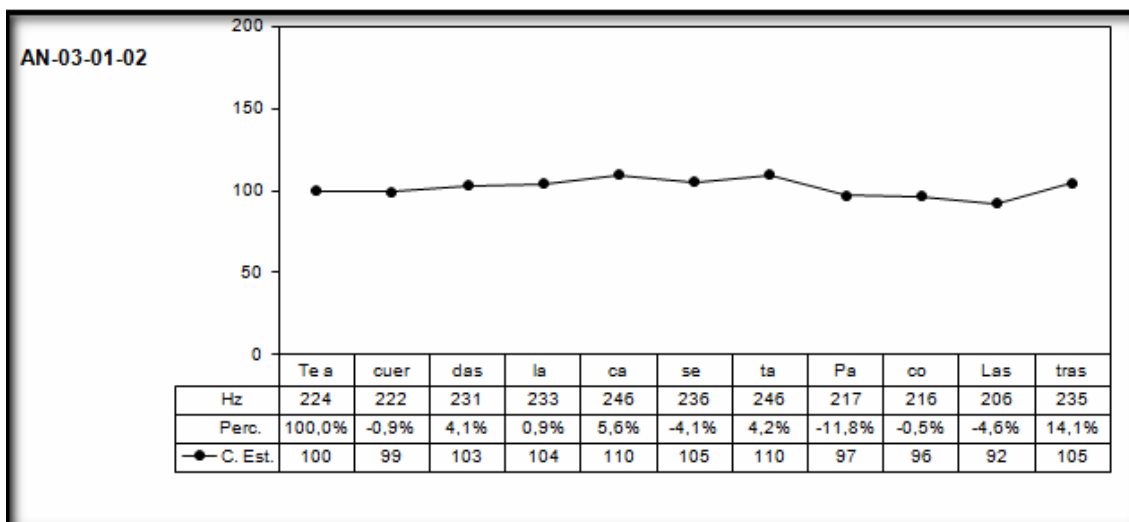


Figura 11 – Curva estándar enunciado AN-03-01-02

El primer gráfico de la figura 11 corresponde a la extracción manual de datos, el segundo a la extracción automática, mediante el programa. Podemos apreciar que en los segmentos tonales “Pa” y “co”, el programa ha extraído dos valores que están dentro de los márgenes del 10% respecto a la extracción manual: 217 y 229-204 en el primer caso y 216 y 201-221 en el segundo. Esta mayor precisión no afecta a la curva melódica que se puede apreciar es idéntica en ambos casos.

El segundo caso –columna 5-, lo comentaremos más adelante, al revisar los motivos de las diferencias superiores al 10% de forma general.

A continuación, verificamos los resultados en función del tipo de informante (masculino / femenino). En las figuras 12 y 13 vemos los resultados.

	Inf	S.Tonales	0%	<2%	< 5 %	5-10 %	>10%	Enunciados 100% OK
VT estándar	F	510	25,27%	79,27%	93,52%	97,62%	2,16%	38
	M	351	37,24%	81,085	87,39%	92,19%	6,91%	15
VT 0.25	F	518	23,33%	74,95%	92,44%	97,41%	2,16%	39
	M	375	36,94%	80,18%	88,89%	95,50%	3,60%	20

Figura 12 –Resultados con la variable género informada

	Inf	3 valores	Valores no informados	Valores extremos	Valores 10%	Valores No 10 %	Valores no script
VT estándar	F	11	2	0	35	14	1
	M	3	9	1	22	4	3
VT 0.25	F	8	0	0	44	11	2
	M	7	0	3	30	14	3

Figura 13 –Alertas con la variable género informada.

Dado que los resultados obtenidos con los dos parámetros, considerados de forma global no ofrecían resultados significativamente diferentes, aunque ligeramente mejores utilizando el parámetro VT no estándar, especialmente en el caso de informantes masculinos, realizamos un análisis cualitativo; revisamos los segmentos del primer pico y el de la inflexión final (IF), desde el último acento, que como indicamos al describir el modelo AMH, es la parte fundamental para la configuración de los diversos patrones. En el primer caso, la bondad de los resultados era cercana al 100% tanto con parámetros estándar como VT con 0,25 y en los casos de discrepancia existe alerta en el fichero correspondiente o bien visual, en la revisión de gráficos.

	IF OK	IF KO	Alertas en IF	OK Con alertas
VT estándar	83,75%	16,25%	6	91,25%
VT 0.25	82,50%	17,50%	10	95%

Figura 14 – Resultados y alertas en IF

En cambio los resultados del segmento tonal de la inflexión final (v. figura 14) presentaban diferencias que nos hicieron decidir por adoptar la solución de no aplicar el valor estándar para el parámetro *voicing threshold*. La única diferencia significativa en la inflexión final se produce en un enunciado (CA-08-02-04), ver más adelante el punto 4.

También el hecho de que al utilizarlo en la fase inicial de otros trabajos, una vez ya validado el script, observamos, haciendo pruebas con ambas versiones, que con

grabaciones de habla espontánea es relativamente frecuente que se produzcan inflexiones, quasi (no)sonidos acompañados de comunicación no verbal, que hacían que se produjera con una frecuencia significativa el fenómeno ya comentado anteriormente, la ausencia de valores de tono (alarma 0), en el caso de utilizar los valores estándar y no reducir el valor de VT.

Como hemos indicado anteriormente, el siguiente paso fue revisar las diferencias superiores al 10% y encontramos la siguiente tipología:

- 1) En primer lugar, una serie de enunciados en los que había problemas de segmentación, es decir que se había hecho de forma incorrecta la primera fase del proceso y de ahí que los datos extraídos por el script sean datos incorrectos. Por ello volvemos a recalcar la importancia del criterio del investigador: segmenta oyendo el enunciado y si percibe un final ascendente, no se puede seleccionar un segmento tonal que tiene rastros de pitch descendente o viceversa.

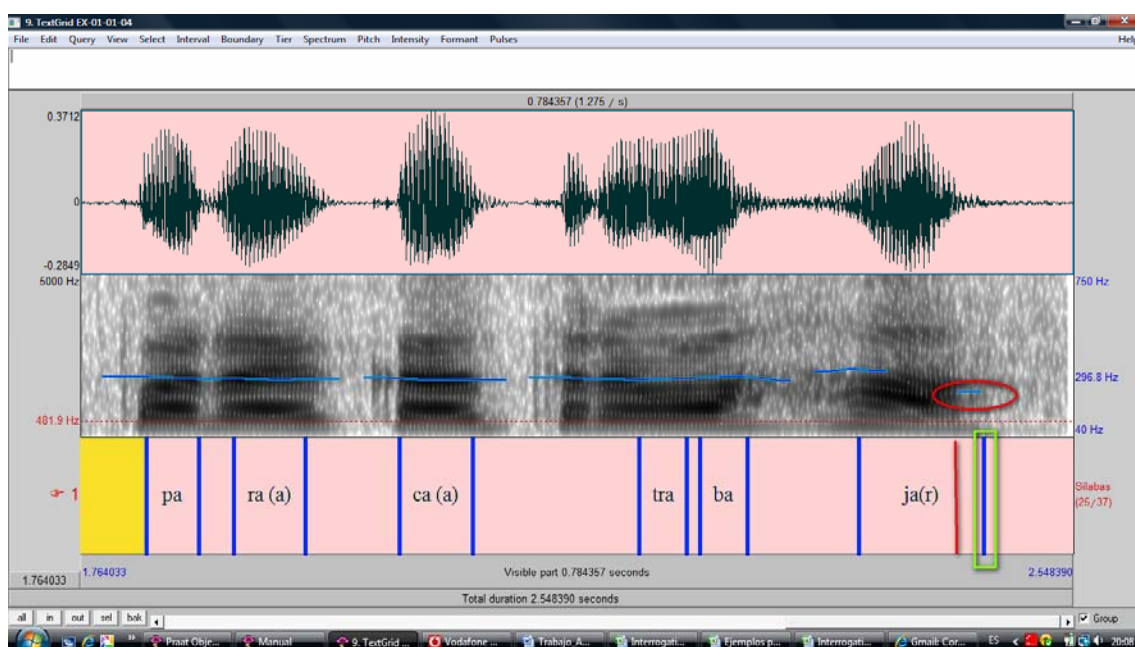


Figura 15 – Textgrid con segmentación errónea

En el enunciado de la figura 15 (hemos ampliado en PRAAT la parte final) se percibe el descenso final cuando lo escuchamos, si la selección la hacemos en el lugar que hemos puesto la barra roja vertical, no se extraerían los datos de pitch marcados con un círculo rojo en el sonograma y los datos posteriores darían lugar a una curva con final ascendente.

Este caso también se da en algunos enunciados a inicio de palabra, son enunciados en que hemos registrado diferencias muy cercanas al 10% (entre el 10,13% y el 12,17%), disminuyendo el segmento 0,05 milisegundos –o menos- los valores son coincidentes con la extracción manual; obviamente por la naturaleza de la diferencia, no afectan a la curva estándar del enunciado. Serían los casos por ejemplo de AN-82-06-01, AN-85-01-07, EX14-01-01, EX21-01-01, PV-01-02-01.

2) 10 enunciados y un total de 16 valores tonales en los que la diferencia está en torno al 50%, es decir, que PRAAT no ha realizado correctamente la extracción; esto también sucede manualmente y el investigador ha de calcularlo atendiendo a los ciclos por segundo del oscilograma. En el caso del análisis manual y según el grado de “entrenamiento” del investigador, se puede detectar en el momento mismo del análisis, sino al realizar la clasificación de los gráficos por patrones. En el caso de la extracción automática, el investigador lo apreciará al comparar los gráficos y ver movimientos no normales, con picos ascendentes o descendentes no naturales, que no encajan en ningún patrón, como podemos apreciar en la figura 16.

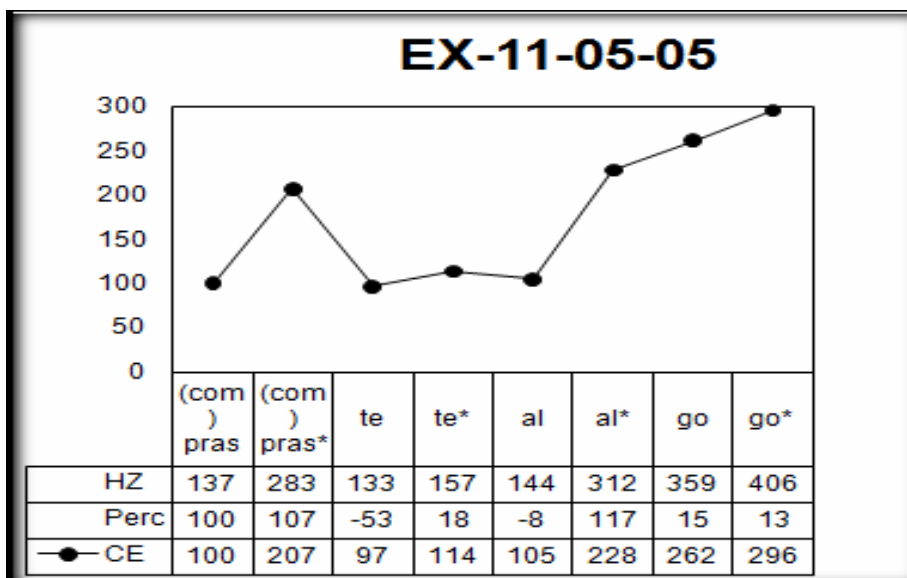
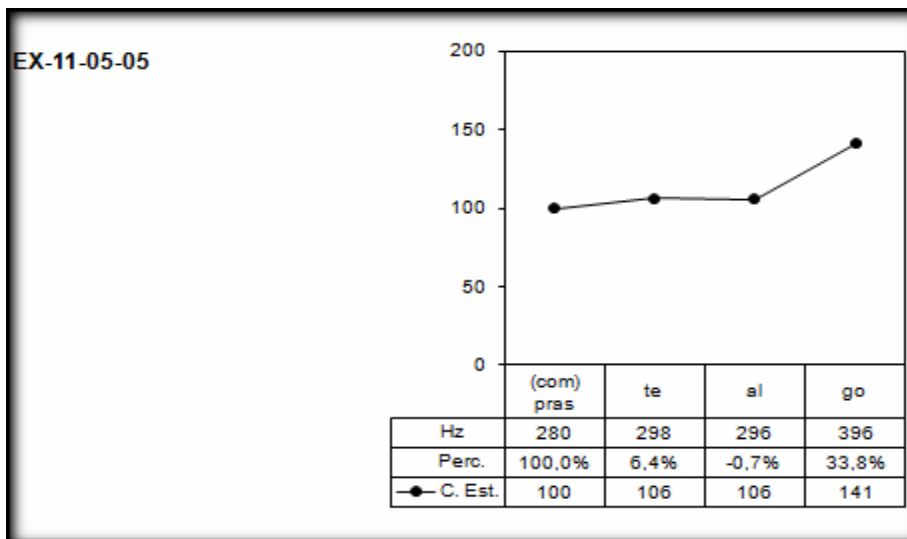


Figura 16 – Curva estándar (manual y script) enunciado EX11-05-05

En el segundo gráfico, obtenido con datos extraídos mediante el script, observamos en los primeros segmentos tonales valores con diferencias en torno al 50% o al 100% :

- (com)pras : 137-283 frente a 280 (51,1%)
- Te / te* : 133-157 frente a 298 (47,3%)
- Al : 144 frente a 296 (48,6%)

La consecuencia son curvas entonativas con inflexiones no reales, que no encajan en ningún patrón. El investigador las deberá detectar en la revisión posterior a la fase de extracción automática de datos, una vez impresos los gráficos.

3) Diferencias superiores al 10% que no responden al caso anterior. Son casos en los que el script detecta valores tonales que en la visualización que efectuamos manualmente no se aprecian, sólo se ha tomado un valor. También se dan casos en que el script captura 3 valores y manualmente sólo se han considerado 2, en este caso existe la correspondiente alarma para la revisión. Esta casuística se da en 9 enunciados, en 6 de ellos no hay impacto en la curva entonativa (cinco casos, AN-26-01-06, AN-63-01-07, CA-43-03-03, CA-40-05-05, MU-01-11-05) o se ha generado una alarma que implicará su revisión manual (un caso, AN-27-02-08).

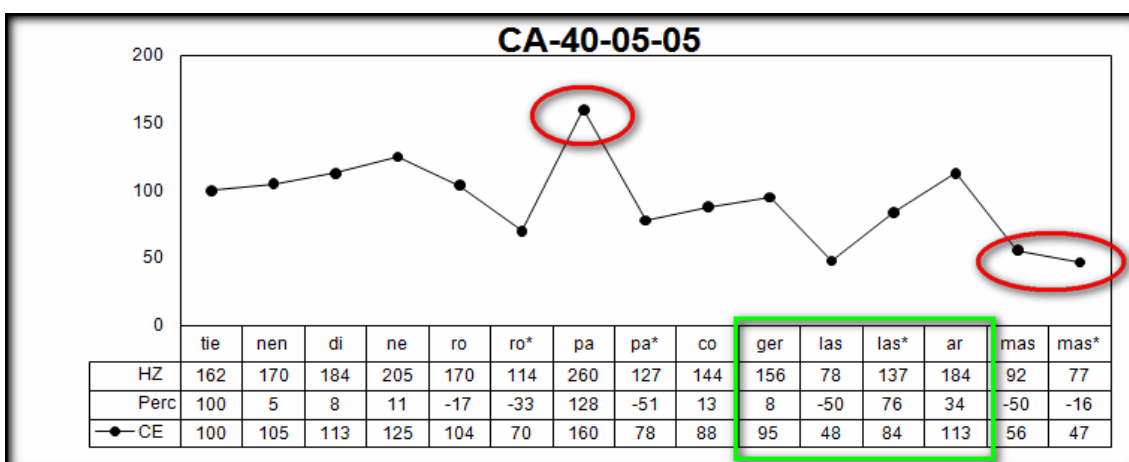
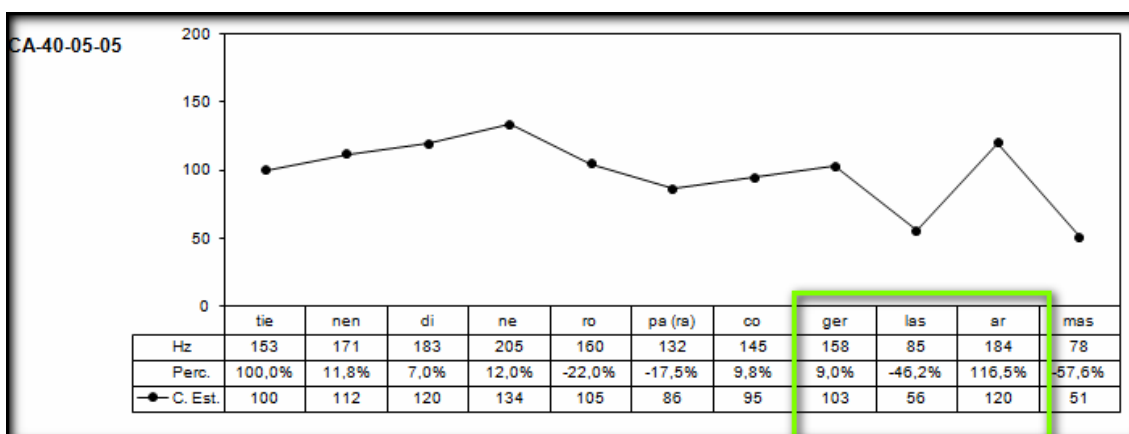


Figura 17 – Curva estándar (manual y script) enunciado CA-40-05-05

En la figura 17, el primer gráfico se ha obtenido con los datos manuales y el segundo a partir de los datos extraídos con el script. Como puede observarse, para el segmento “las”, el script proporciona dos valores pero la curva del gráfico no se ve afectada, el segundo valor está en el camino hacia los 184 hercios del siguiente segmento –“ar”-. Hemos marcado en rojo, dos fenómenos que presenta este enunciado: por una parte un valor para el que ha calculado el doble, “pa”, y por otra parte, en el segmento final, dos valores, aunque el porcentaje desde la última átona es equivalente.

En los tres enunciados restantes, en los que la curva estándar resultante no es correcta, en dos de ellos el error está en la primera sílaba, se han detectado dos valores, que en la anotación manual no se han recogido, son los enunciados CM-18-02-06 (v. figura 18) y MU-76-08-02.

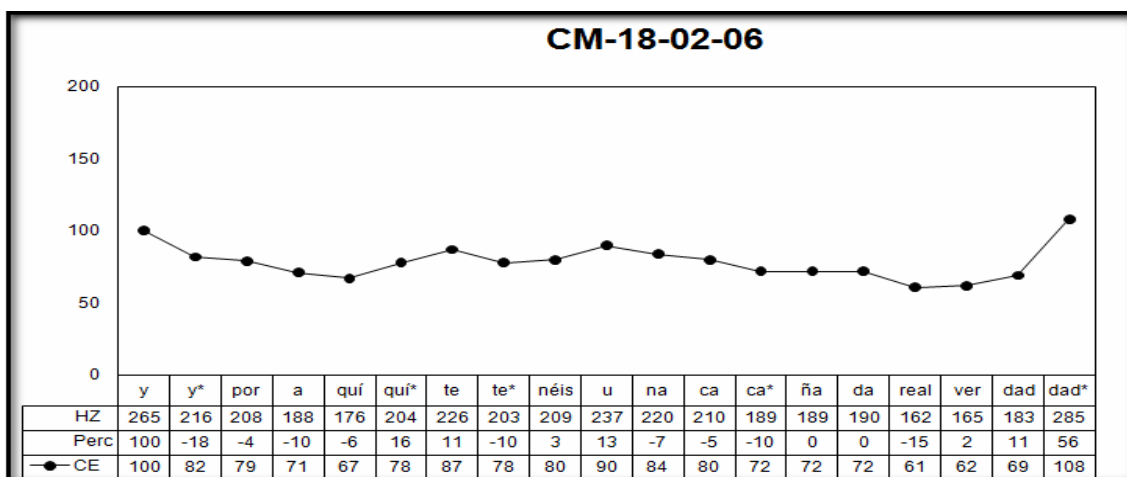
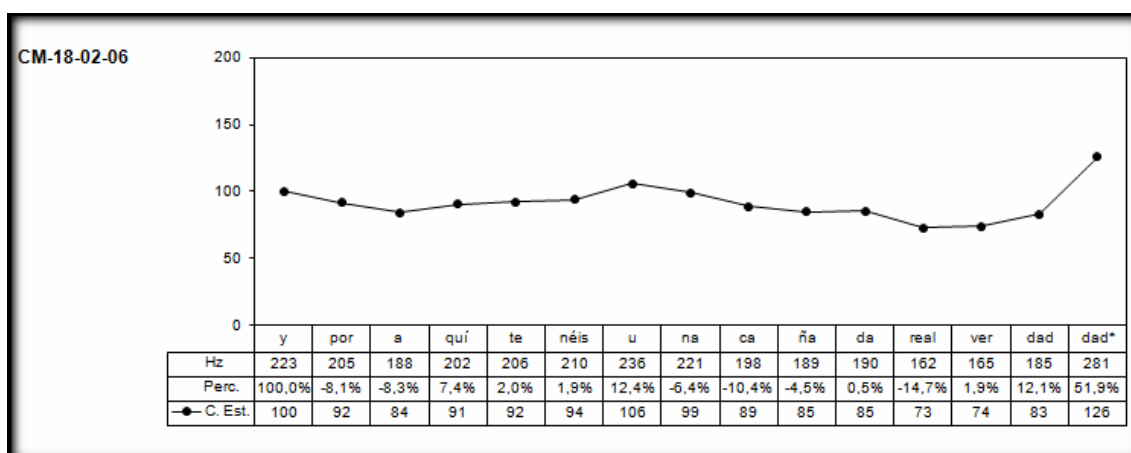


Figura 18 – Curva estándar (manual y script) enunciado CM-18-02-06

Creemos que estos datos no son significativos, podríamos considerarlo una variación micromelódica o un rasgo de énfasis inicial –relacionado seguramente con la duración de la vocal- que utilizando el parámetro *voicing thresold* se hace patente, la percepción de hecho es la de una vocal más larga.

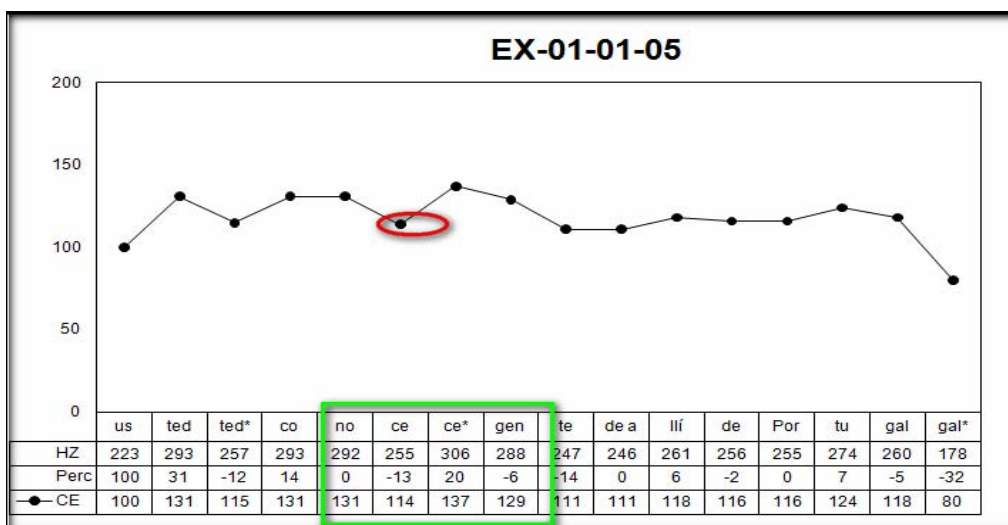
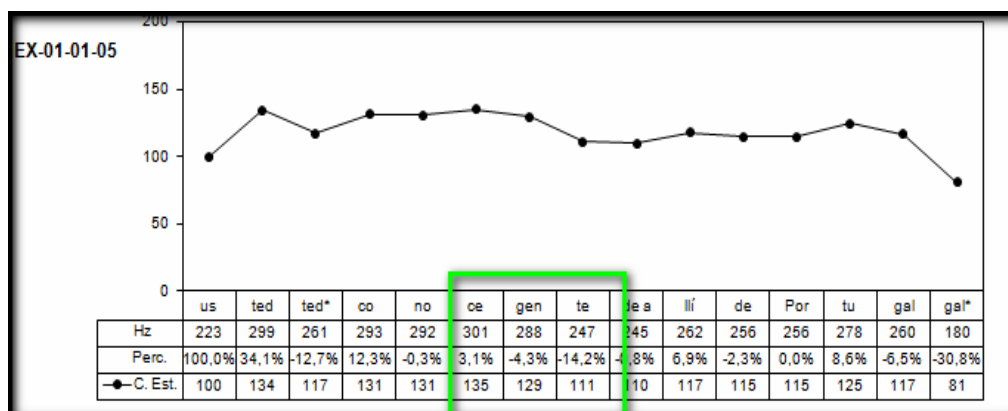


Figura 19 – Curva estándar (manual y script) enunciado EX01-0105

En el otro enunciado (figura 19), en la extracción automática observamos que se detecta un segundo segmento tonal en ‘ce’, pero que hace que el cuerpo de la curva entonativa baje inicialmente. Este hecho se detectará en la revisión visual de los gráficos, al realizar la clasificación y el investigador deberá valorar si el hecho es relevante o no, de acuerdo a los patrones definidos.

- 4) Diferencias superiores al 10%, con inflexión final diferente en la medición manual y la realizada mediante programa, se produce únicamente un caso, el enunciado CA-08-02-0 (v. figura 20), como en los casos anteriores, el primero obtenido manualmente y el segundo con los datos extraídos por programa.

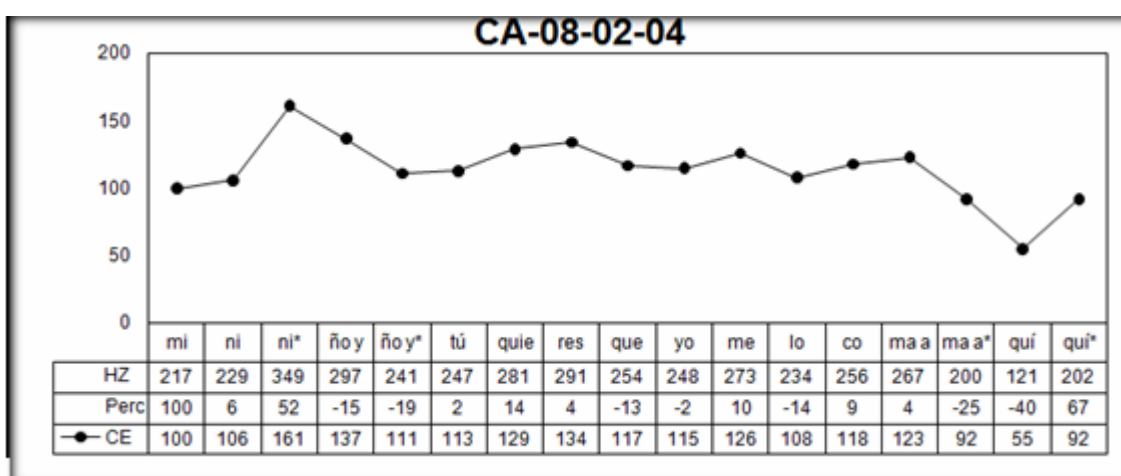
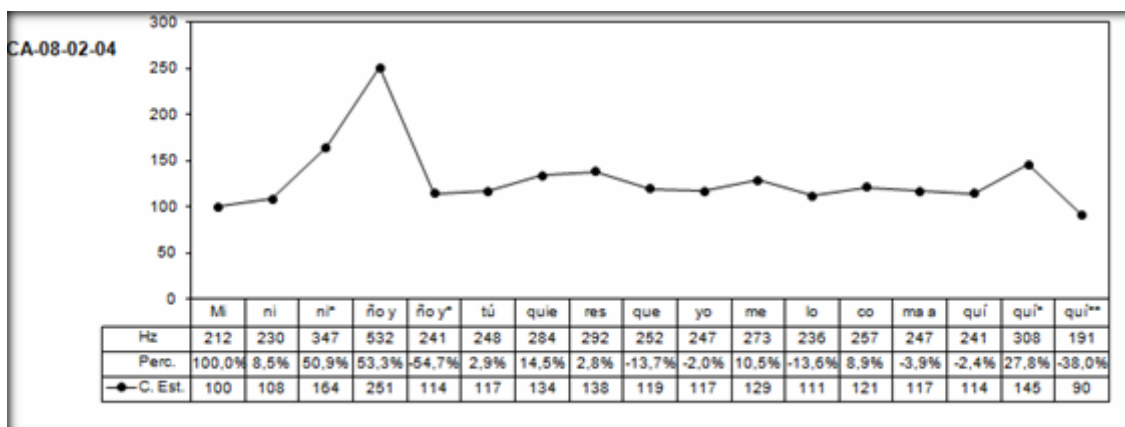


Figura 20 – Curva estándar (manual y script) enunciado CA-08-02-04

Es éste el único caso que depende totalmente de la percepción del investigador ya que el resultado que obtiene PRAAT mediante el programa encaja en un patrón melódico.

6. Protocolo para extracción semiautomática de datos tonales de un enunciado

A continuación detallamos los pasos a seguir para facilitar la utilización a usuarios con distinto nivel de familiarización y destreza en el uso de PRAAT, siguiendo las etapas establecidas en el apartado metodológico:

a) Segmentación y etiquetaje de los segmentos tonales.

En primer lugar, crearemos el fichero *TextGrid* del sonido a analizar:

- Seleccionamos el fichero con el que vamos a trabajar, previamente incorporado desde nuestro repositorio con la instrucción “Read” de la barra de herramientas y pulsaremos el botón “Anotate” (figura 21).

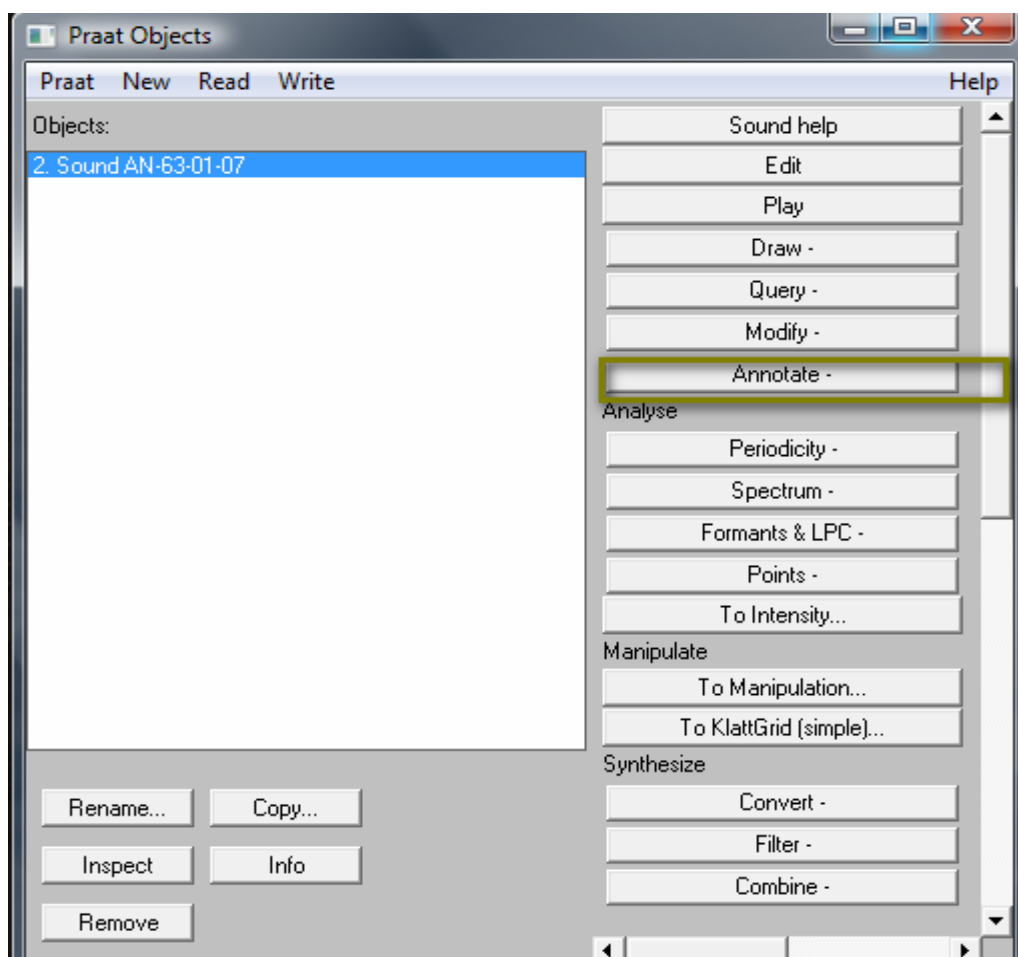


Figura 21 – Creación textgrid, paso 1

- En el desplegable, seleccionamos la opción “To text Grid” (figura 22)

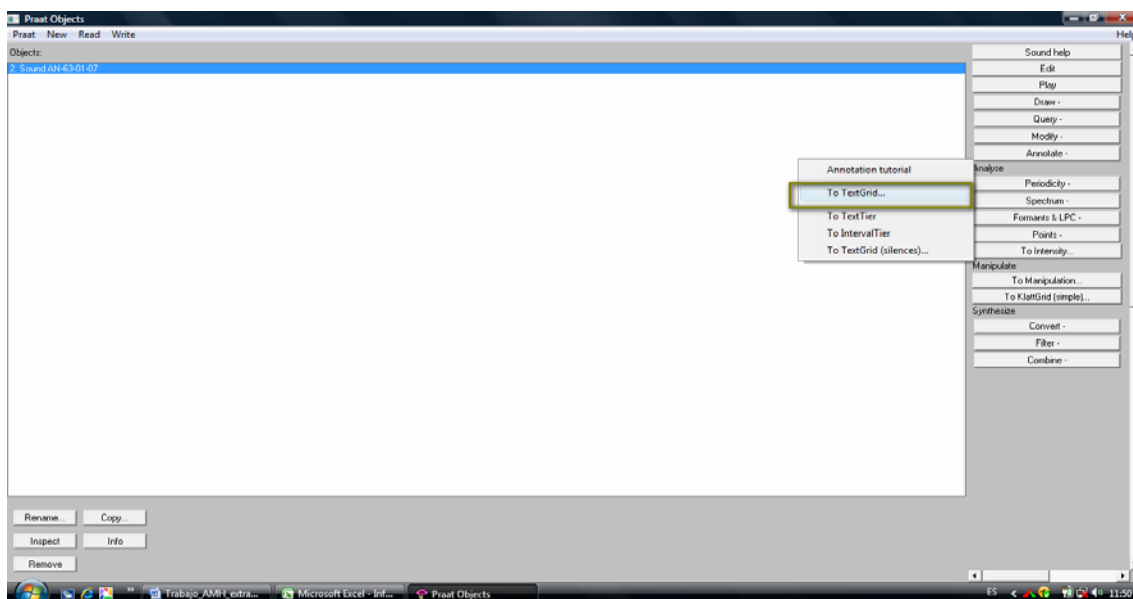


Figura 22 – Creación textgrid, paso 2

- Esta opción nos presenta la siguiente pantalla, en la que indicamos un nombre para el contenido de la información que recogeremos en el “tier”, nivel, que vamos a crear en el *TextGrid*, en nuestro caso S.Tonal. Es importante indicar que el nombre ha de ser bien una palabra, bien dos separadas por un carácter, si dejamos un blanco entre las palabras, PRAAT crea tantos “tier” como palabras introduzcamos.

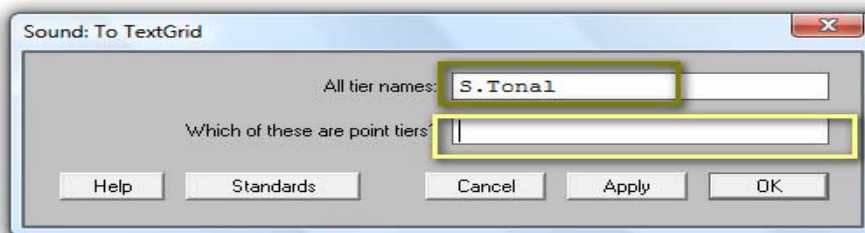


Figura 23– Creación textgid, paso 3

- Tras pulsar el botón “OK”, en nuestra área de trabajo nos aparecerá además del sonido el fichero *TextGrid* (figura 24).

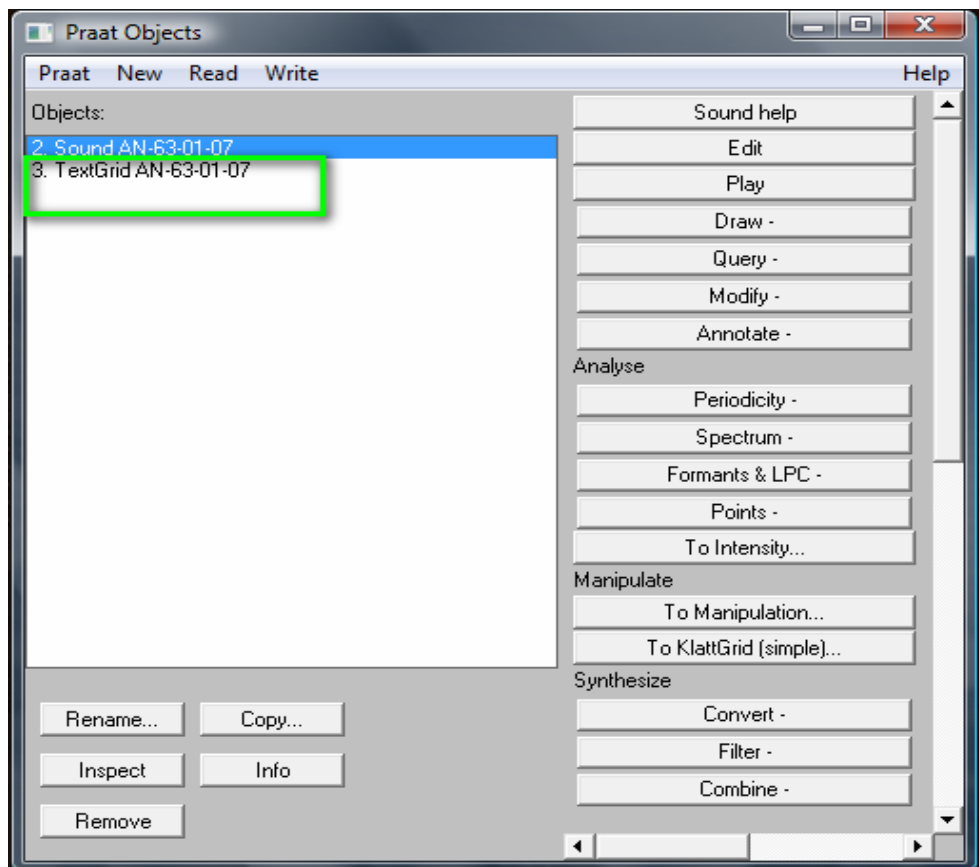


Figura 24– Creación textgid, paso 4

- A continuación, seleccionaremos los dos elementos, sonido y *TextGrid* y pulsaremos el botón “EDIT” con lo cual obtendremos la siguiente pantalla :

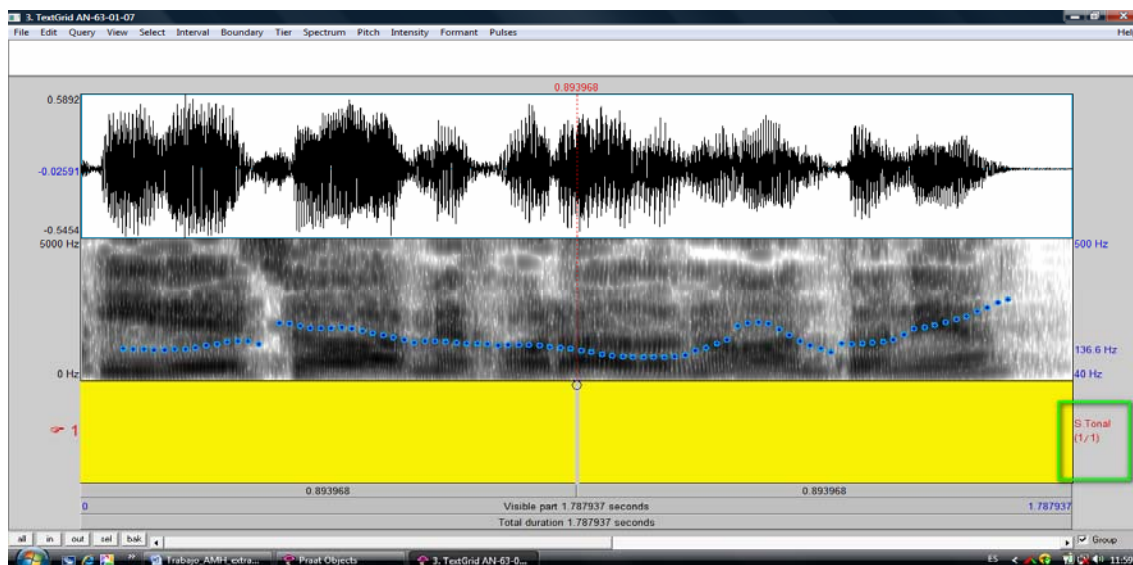


Figura 25– Creación textgid, paso5

En la figura 25, podemos observar, de amarillo, el área previamente definida (S. Tonal), en la que iremos marcando los distintos segmentos tonales del enunciado analizado. En esta pantalla la primera vez que iniciemos el proceso de análisis (no para cada enunciado), revisaremos los valores de configuración de la obtención del pitch. Esta acción se realiza en el menú *pitch* (figura 26).

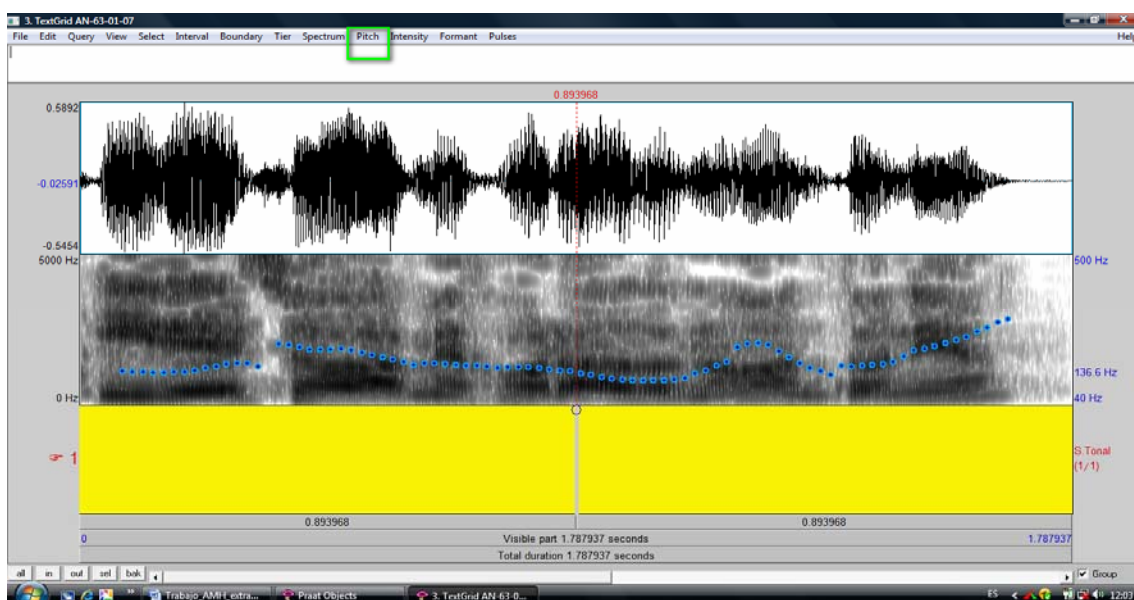


Figura 26– Creación textgid, paso 6

Revisaremos tanto los parámetros de configuración (“*pitch settings*”) como los avanzados (figura 27).

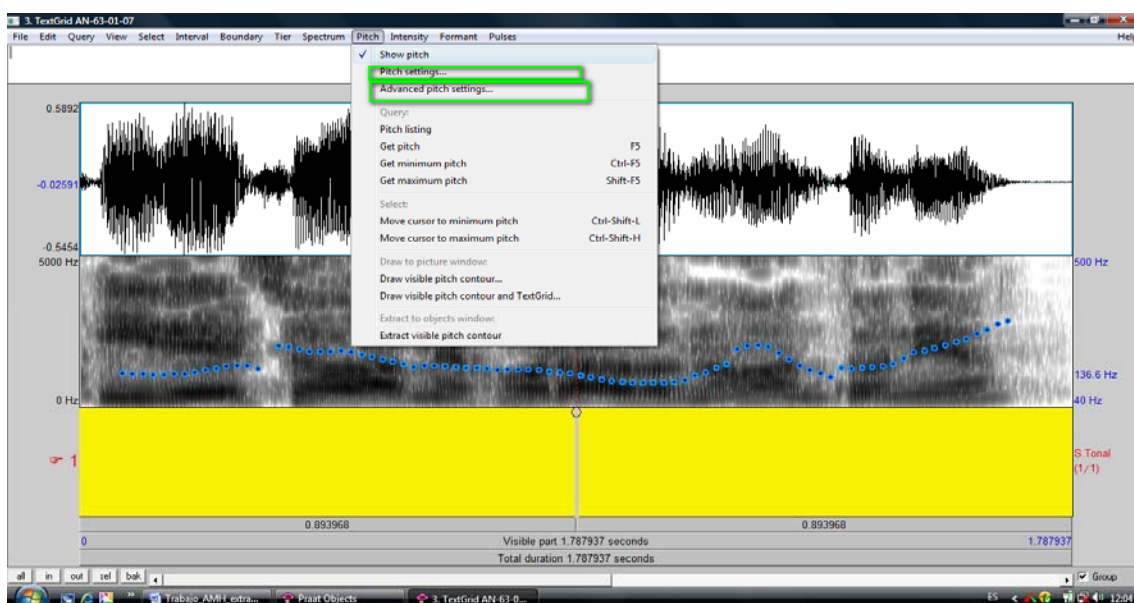


Figura 27– Creación textgid, paso 7

En la primera opción (*Pitch settings*), como vemos en la figura 28

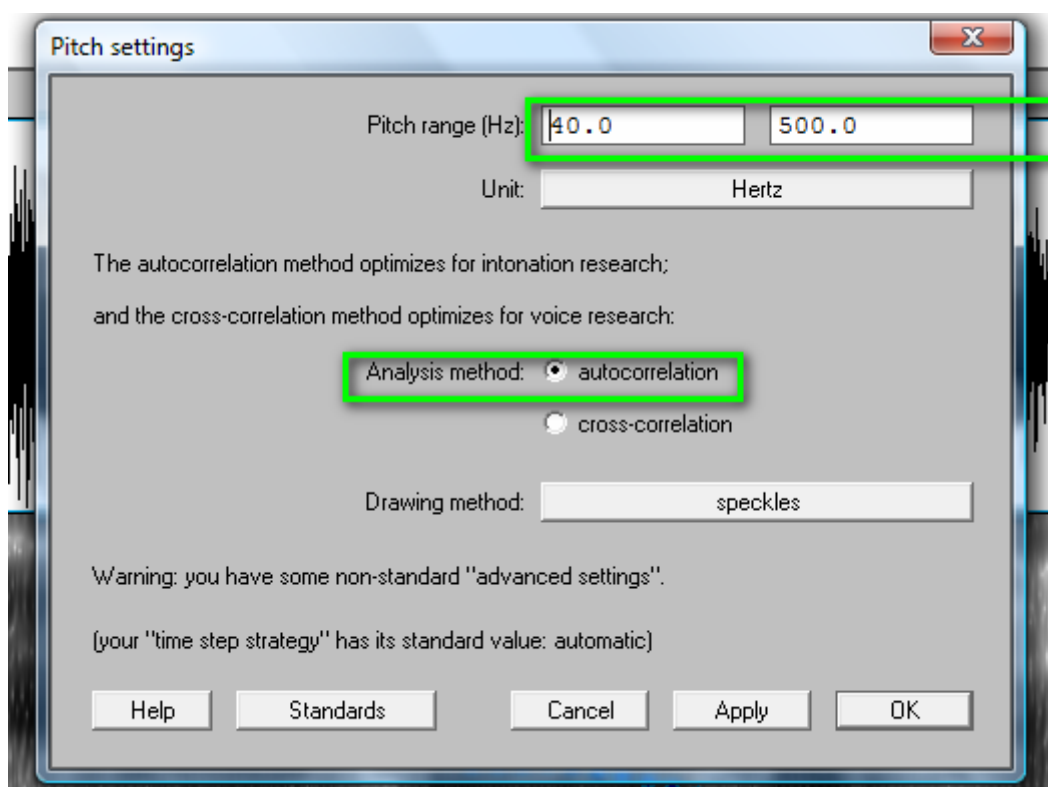


Figura 28– Creación textgid, paso 8

Indicaremos el rango estándar de extracción, como decíamos en el apartado 6, el investigador modificará este rango si durante el proceso de etiquetaje es necesario para la correcta identificación de los segmentos tonales; los cambios más habituales (que no frecuentes) son del límite superior, bajándolo para voces muy graves (a 350 Hz, por ejemplo) y subiéndolo para voces muy agudas

Como ya indica la propia pantalla el método de análisis que debemos seleccionar es el de “*autocorrelation*”, que es el óptimo para investigar la entonación.

En la pantalla de parámetros avanzados (figura 29) modificaremos dos valores respecto a los estándar para igualarlos con los que utiliza el programa. Como indicamos al describir el programa, el primero (*very accurate*) afecta a la configuración de la pantalla y el segundo (*voicing treshold*) al umbral de consideración de valores tonales o no. Dado que trabajamos con habla espontánea, no de laboratorio, sobre todo en los inicios y final de enunciado, con el valor estándar el algoritmo de detección de “pitch” no lo detecta; los resultados obtenidos vimos que fueron satisfactorios. Si no los cambiamos, obviamente, no hay impacto en los resultados de la extracción, pero la experiencia es que ayudan en el proceso de identificación y segmentación.

Pulsaremos OK para finalizar los cambios en ambas opciones.

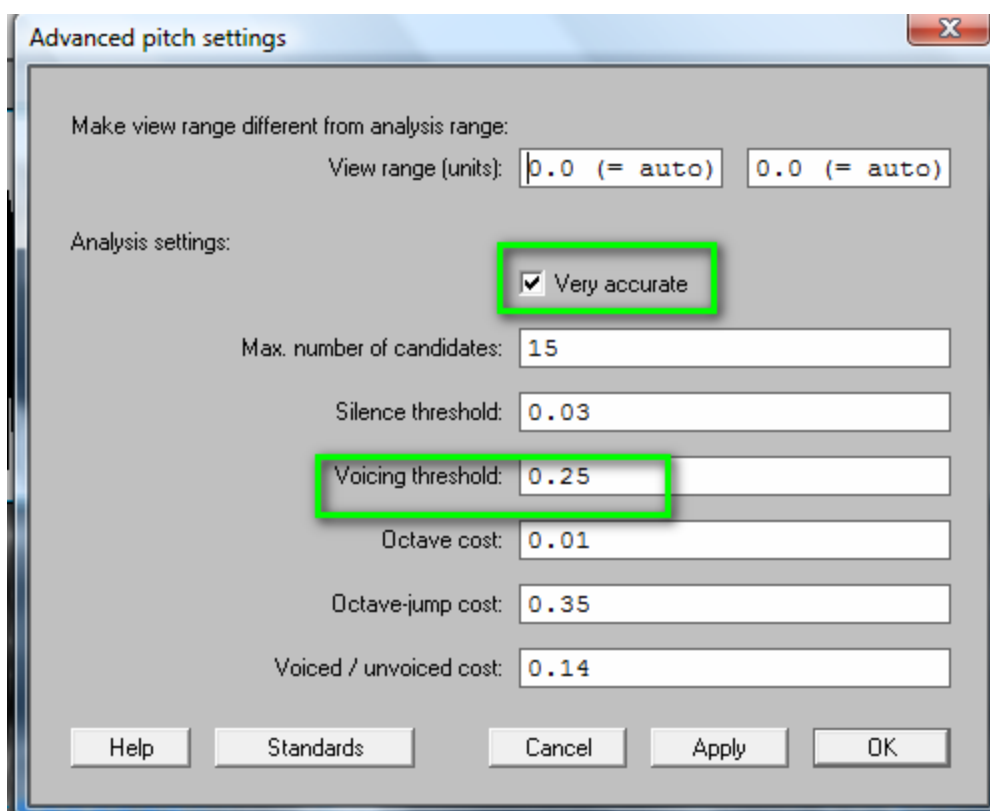


Figura 29– Creación textgid, paso 9

Los siguientes pasos, son los ya conocidos y descritos en el protocolo de análisis melódico, identificar las vocales (y laterales o nasales en los casos que sean pertinentes, en lugar de tomar los datos manualmente, las identificaremos y marcaremos para su posterior análisis automático. Seleccionamos con el cursor, tal y como podemos ver en la figura 30.

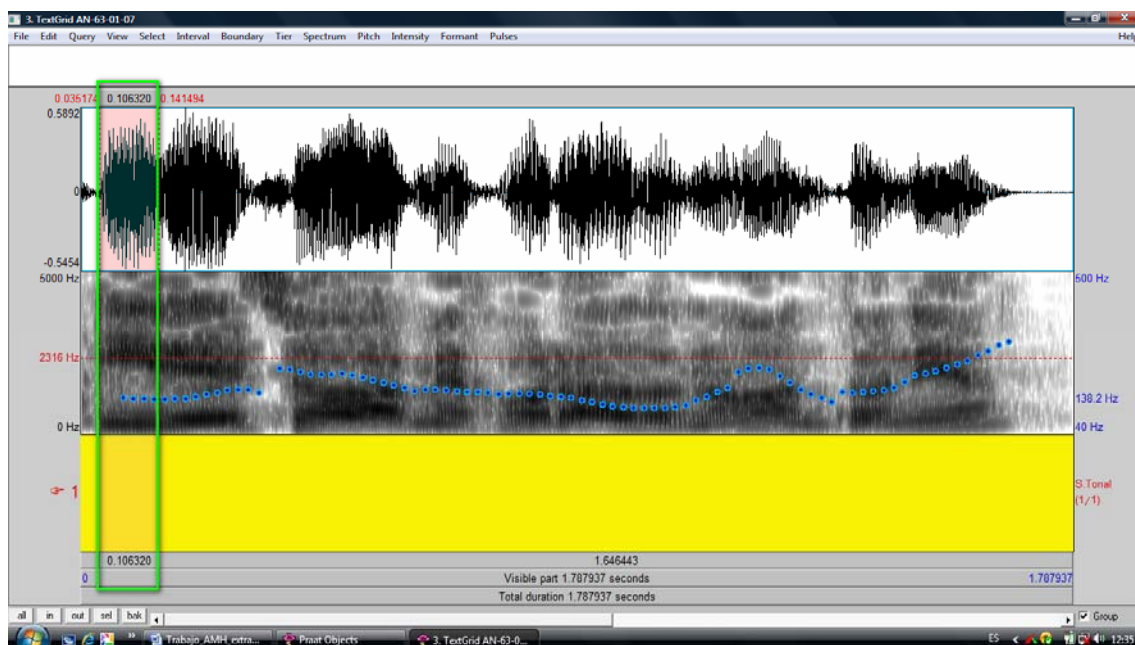


Figura 30– Creación textgrid : selección segmento tonal

En el ejemplo tenemos seleccionado mediante los cursores de PRAAT el fragmento “Y”, una vez seleccionado, para marcarlo (y posteriormente analizarlo), pondremos las marcas, se puede hacer mediante el menú “Tier” o “Interval”, pero lo más práctico es con las teclas de función que ya tiene predefinidas PRAAT : CTRL+F1.

Se puede observar que en la zona izquierda no hay indicador de “pitch” (línea azul) pero sí que hay marcas de los formantes y si seleccionamos sólo la zona sin información suena ‘i’, la percibimos. Reiteramos lo ya apuntado, el investigador debe aportar su criterio y seleccionar los fragmentos vocálicos que ve y oye; el algoritmo no muestra los valores en el 100% de los casos.

En el área de texto superior, ponemos la información del segmento tonal en cuestión en cuestión, apareciendo entonces informada en el *textgrid*.

Una vez identificadas todas las vocales o grupos a analizar, los etiquetamos adecuadamente, únicamente con caracteres alfabéticos (o fonéticos), sin puntos, comas, admiraciones u otros elementos gráficos, del discurso escrito, tal como podemos ver en las figura 31 y 32.

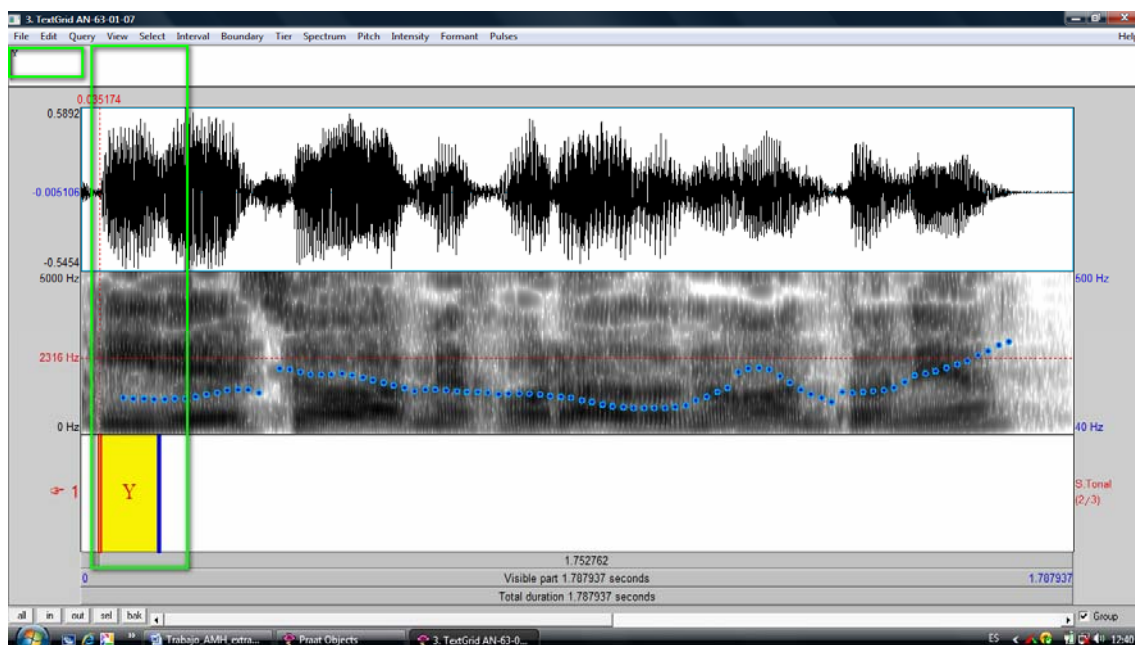


Figura 31– Creación textgrid :identificación segmento tonal

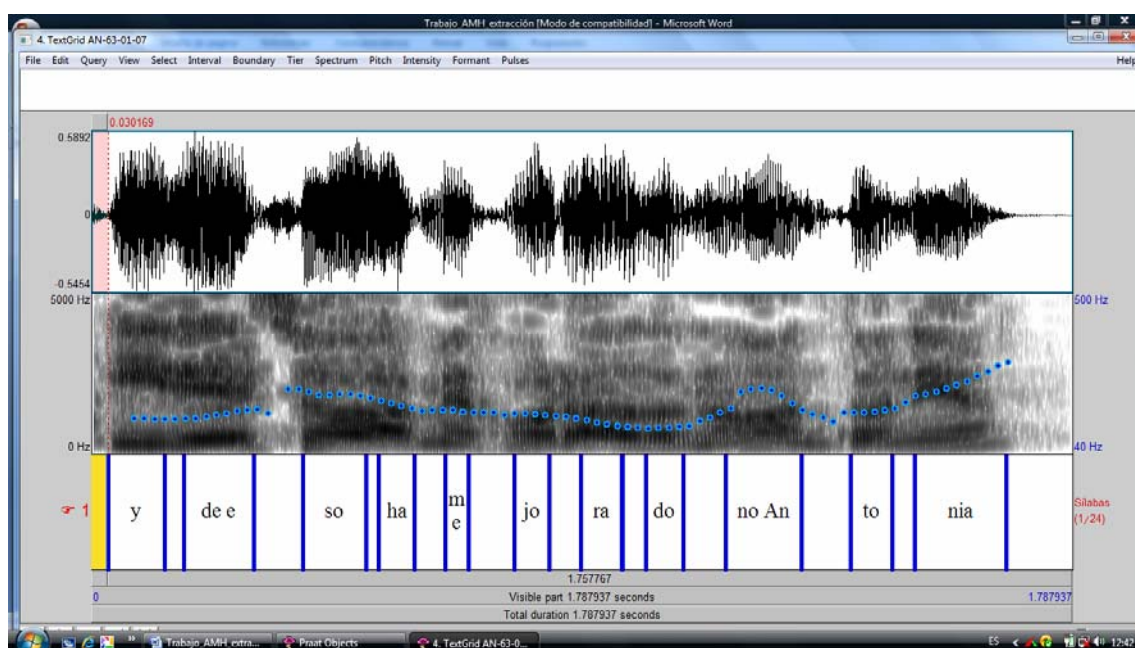


Figura 32– Creación textgrid :identificación de todos los segmentos tonales

En el primer segmento, en la figura 33 en amarillo, indicaremos si la voz es masculina (m) o femenina (f), para poder aplicar un algoritmo diferente cuando pasemos el programa de extracción de datos, como se puede ver, esta imagen es una selección ampliada del sonograma, que no afecta a los datos.

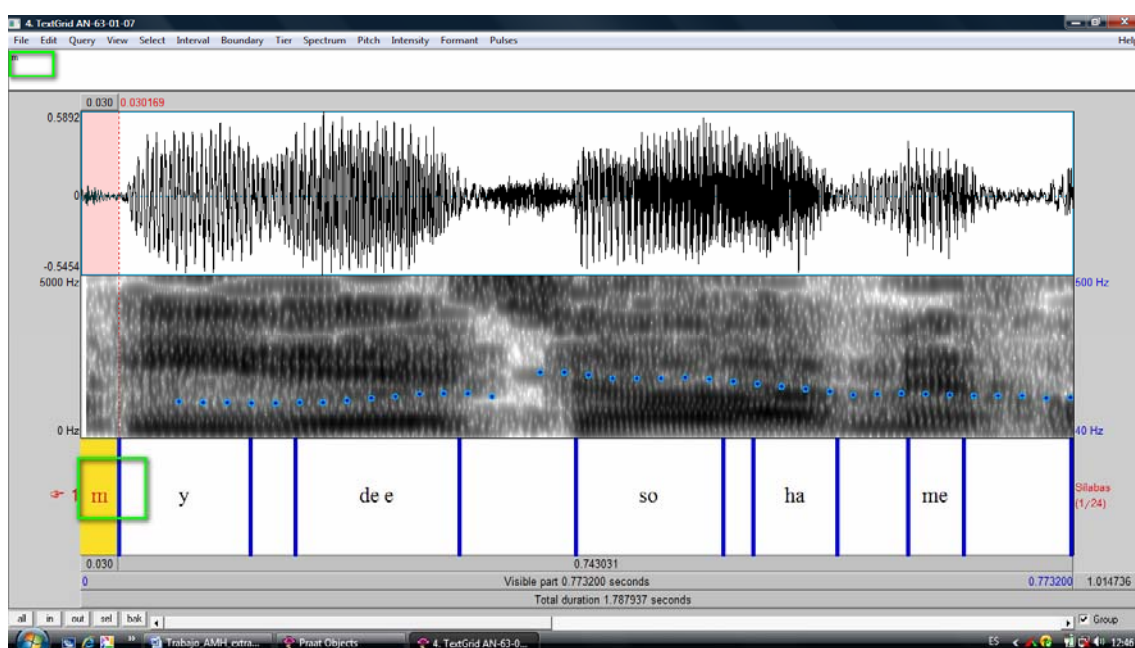


Figura 33– Creación textgrid :identificación tipo de voz

Una última observación importante respecto a la selección y etiquetaje: las líneas azules forman parte de la selección, se analizan.

Concluido el proceso de etiquetaje, cerramos el *textgrid*, con la opción *Close*, del menú (figura 34).

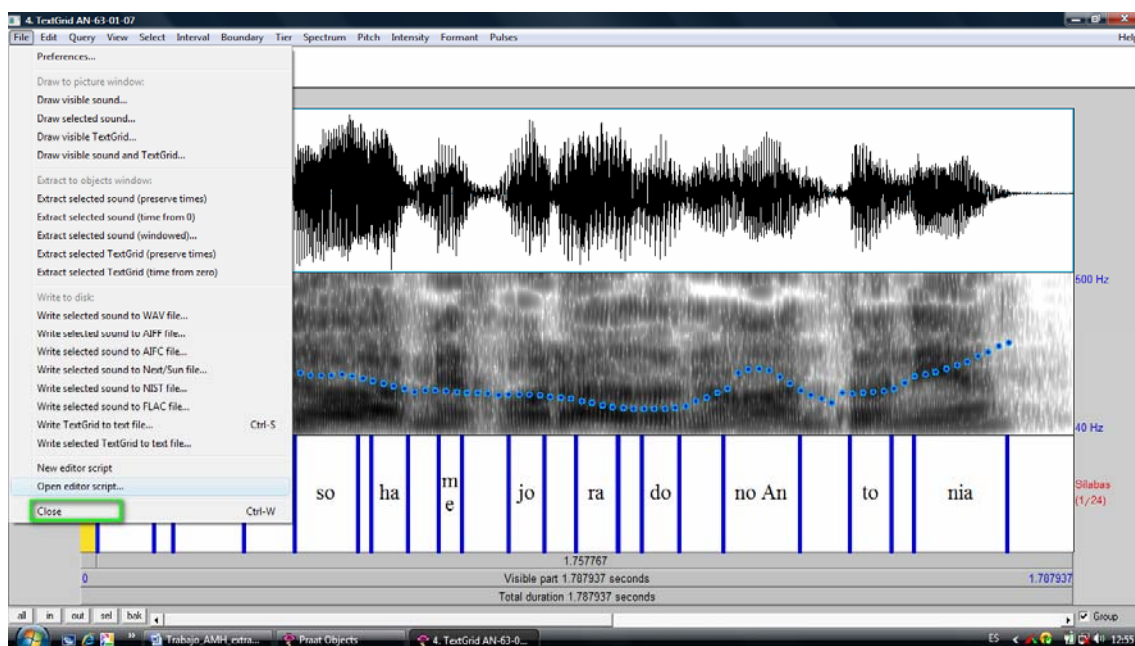


Figura 34– Creación textgrid

- Finalmente, salvaremos el fichero de *TextGrid*, que pasa a formar parte del corpus junto con el fichero del sonido propiamente dicho. Seleccionamos únicamente el fichero de *TextGrid* de nuestra área de trabajo y lo grabamos con la opción *Write* de la barra de menú, en el directorio que consideremos oportuno.

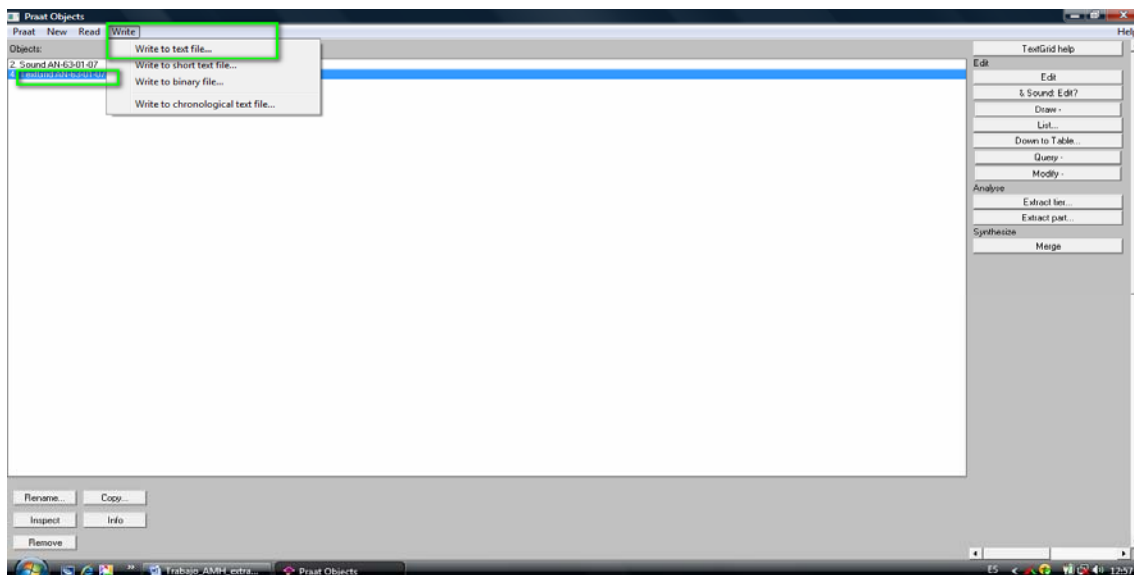


Figura 35– Creación textgrid . grabación del fichero

Nota: cada vez que realicemos un cambio en el *textgrid*, debemos recordar salvarlo; en caso contrario perderemos las modificaciones para posibles usos posteriores.

b) Extracción de datos.

Antes de realizar este paso habremos instalado los programas⁷.

Al seleccionar esta opción (AMH-Extracción F0) se realiza el proceso de extracción de datos y se generan dos ficheros: uno con los datos tonales (F0) y, si es el caso, un fichero con alertas (v. figura 36)

⁷ Ver artículo “Scripts en PRAAT para la extracción de datos tonales y curva estándar” en este mismo volumen

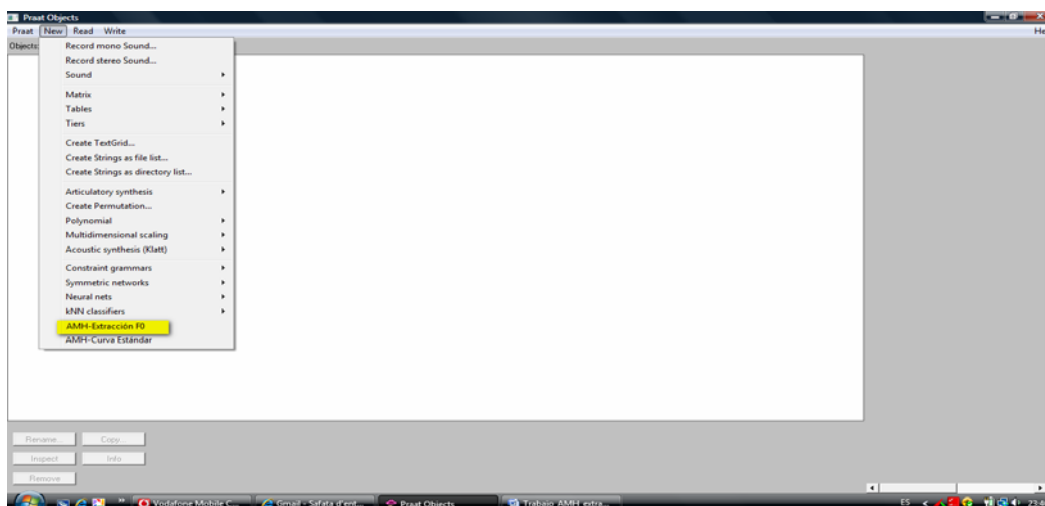


Figura 36– Proceso script 1 : extracción F0.

Si en el directorio dónde está nuestro corpus falta algún sonido o textgrid o bien la “pareja” sonido-textgrid no tiene exactamente el mismo nombre el proceso se detiene con un error como el que se muestra en la figura 37

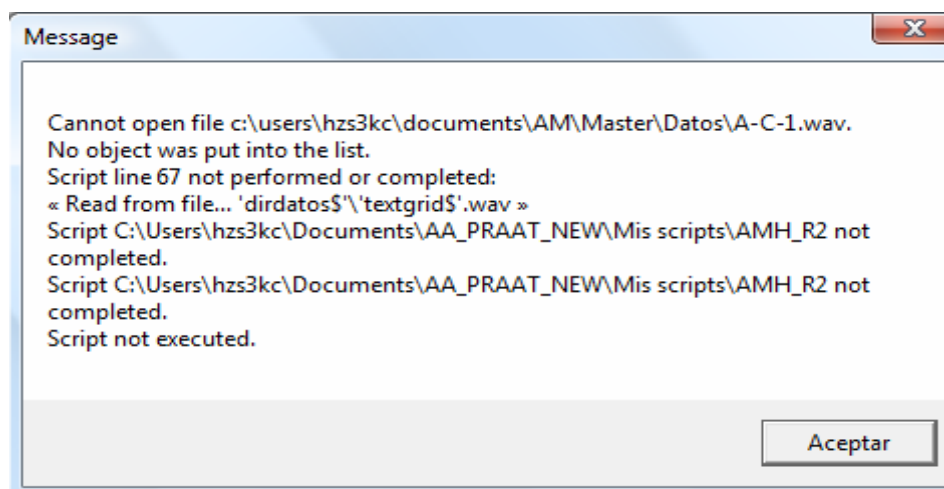


Figura 37– Ejemplo mensaje error textgrid

c) Revisión de alertas.

En el directorio “C:\..... \xxxxx”, según hayamos definido en el proceso de instalación del programa, tendremos la siguiente estructura de carpetas (v. figura 38)

Nombre	Fecha modificación	Tipo	Tam
ficheros_F0_Revisar	18/08/2010 13:57	Carpeta de archivos	
ficheros_F0	18/08/2010 13:57	Carpeta de archivos	
CURVA_F0	18/08/2010 13:57	Carpeta de archivos	

Figura 38– Carpetas de trabajo del textgrid

En la subrayada, ficheros_f0_Revisar, estarán todos los ficheros que tengan alguna alerta que revisar, veámoslo, en la figura 39, con un ejemplo de nuestra investigación

Nombre	Fecha modificación	Tipo	Tamaño	Etiquetas
Ai-23	13/08/2010 19:50	Documento de tex...	1 KB	
Ai-32	13/08/2010 10:39	Documento de tex...	1 KB	
AN-27-02-08	13/08/2010 11:19	Documento de tex...	1 KB	
AN-63-01-07	13/08/2010 11:20	Documento de tex...	1 KB	
AN-71-01-07	13/08/2010 11:20	Documento de tex...	1 KB	

Figura 39– Relación de ficheros de alertas

Si abrimos uno de los ficheros de alerta, podemos ver el segmento tonal afectado y la posible incidencia, el programa ha extraído tres valores para dicho segmento, como se puede observar en la figura 40.

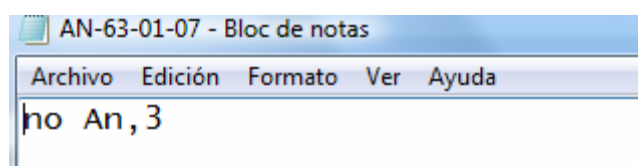


Figura 40– Contenido de un fichero de alertas

lo podemos comprobar si vamos a la carpeta 'ficheros_F0' y lo consultamos (v. figura 41)

```

AN-63-01-07 - Bloc de notas
Archivo Edición Formato Ver Ayuda
Segmento, HZ, Perc, CE
y, 138, 0, 0
de e, 138, 0, 0
de e*, 164, 0, 0
so, 208, 0, 0
ha, 192, 0, 0
ha*, 166, 0, 0
me, 161, 0, 0
jo, 152, 0, 0
ra, 137, 0, 0
ra*, 117, 0, 0
do, 113, 0, 0
no An, 162, 0, 0
no An*, 225, 0, 0
no An**, 167, 0, 0
to, 150, 0, 0
nia, 200, 0, 0
nia*, 296, 0, 0

```

Figura 41– Contenido de un fichero de F0

Veamos ejemplos de cómo están informadas las otras alertas definidas, en el primer caso (figura 42) PRAAT no ha podido extraer datos tonales del segmento “cues” de este enunciado:

```

NA-02-05-04 - Bloc de notas
Archivo Edición Formato Ver Ayuda
cues, 0

```

Figura 42– Contenido de un fichero de alertas

En este caso, en el fichero “F0”, con los datos tonales, encontraremos grabado un 22, esto es debido a que el lenguaje PRAAT da un error al intentar calcular la curva estándar cuando un valor está a 0 y no se ha modificado con el valor correcto (v. figura 43).

```

NA-02-05-04 - Bloc de notas
Archivo Edición Formato Ver
Segmento, HZ, Perc, CE
yo, 104, 0, 0
me, 103, 0, 0
sien, 112, 0, 0
sien*, 98, 0, 0
to, 109, 0, 0
to*, 88, 0, 0
más, 97, 0, 0
sa, 95, 0, 0
tis, 96, 0, 0
fe, 104, 0, 0
cho, 112, 0, 0
cho*, 96, 0, 0
de, 91, 0, 0
de*, 81, 0, 0
cues, 22, 0, 0
tío, 96, 0, 0
nes, 96, 0, 0
nes*, 83, 0, 0
dis, 97, 0, 0
tin, 101, 0, 0
tas, 96, 0, 0
a e, 88, 0, 0
sa, 90, 0, 0
sa*, 331, 0, 0

```

Figura 43– Contenido de un fichero F0

d) Corrección de datos.

El investigador revisará el sonograma para comprobar si los datos son correctos.

```

AN-63-01-07 - Bloc de notas
Archivo Edición Formato Ver Ayuda
Segmento, HZ, Perc, CE
y, 138, 0, 0
de e, 138, 0, 0
de e*, 164, 0, 0
so, 208, 0, 0
ha, 192, 0, 0
ha*, 166, 0, 0
me, 161, 0, 0
jo, 152, 0, 0
ra, 137, 0, 0
ra*, 117, 0, 0
do, 113, 0, 0
no An, 162, 0, 0
no An*, 225, 0, 0
no An**, 320, 0, 0
to, 159, 0, 0
nia, 200, 0, 0
nia*, 296, 0, 0

```

Figura 44– Contenido de un fichero F0 corregido manualmente

Si no lo son, modificará manualmente el fichero y lo salvará, supongamos, siguiendo con el ejemplo anterior que el último valor es incorrecto y que realmente es de 320 Hz como se observa en la figura 44.

e) *Elaboración curva estándar*

Una vez corregidos y salvados todos los ficheros con alertas, pasamos a calcular la curva estándar, para ello, en el menú “New” de la pantalla “PraatObjects”, seleccionaremos la opción Curva Estándar, como indicamos en la figura 45

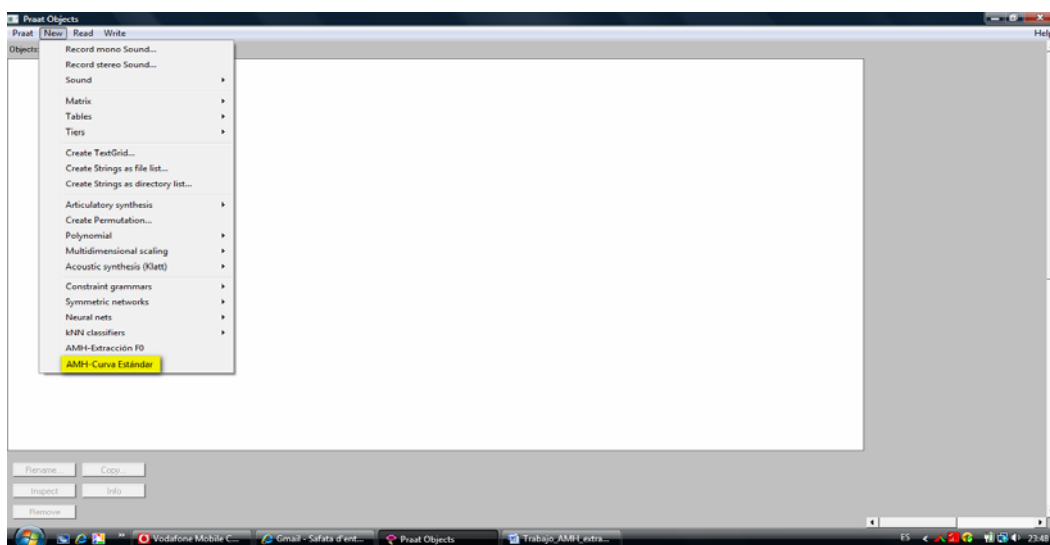


Figura 45 – Proceso script 2 : cálculo curva estándar.

que obtendrá los datos para todos los enunciados que tenemos en el directorio F0, resultado del primer script, con las modificaciones manuales pertinentes.

Como hemos indicado anteriormente, el proceso está concebido por pasos: una vez realizada la extracción de los datos tonales de un grupo de enunciados –una variedad dialectal por ejemplo- (b), revisamos todas las alertas de ese grupo de enunciados (c), las comprobamos y realizamos las correcciones que procedan y, finalmente obtenemos la curva estándar (d y e).

f) *Generación de los gráficos con la curva estándar.*

Una vez tenemos todos los ficheros con los datos de curva estándar correctamente calculados ya podemos realizar los gráficos que nos permiten visualizar los contornos. Este proceso se puede realizar mediante una hoja Excel, ya que este programa dispone de una buena herramienta de generación de gráficos.

7. Conclusiones

Si repasamos los objetivos que nos planteábamos al iniciar nuestra investigación, creemos que se han cumplido de forma satisfactoria: en este trabajo hemos presentado un programa que permite realizar la extracción (semi)automática de datos tonales; hemos verificado la validez de los datos que proporciona de forma experimental y finalmente, hemos desarrollado un detallado protocolo de utilización que consideramos que permitirá su uso a los investigadores, independientemente de su conocimiento previo de PRAAT.

Creemos, por otra parte, que este trabajo puede tener continuidad en el desarrollo de nuevas “herramientas”, nuevos programas que permitan que el esfuerzo principal de los investigadores, nuestro esfuerzo, nuestra labor, se centre no tanto en la obtención de datos sino en su análisis e interpretación. Por ello, en paralelo al estudio de la entonación y sus implicaciones y aplicaciones didácticas, tenemos identificadas algunas funcionalidades a desarrollar: confección automática de gráficos, simplificación de la intervención del investigador en la obtención de los ficheros de curva estándar y en el reproceso de los enunciados con alertas, parametrizar el porcentaje significativo de variación tonal ya que estudios en curso parecen indicar que en el caso del portugués, este valor es inferior al 10% u otras modificaciones derivadas del avance en la aplicación del *Análisis Melódico del Habla*, como pudiera ser, por ejemplo, el análisis del último segmento tonal.

8. Referencias bibliográficas

BORESMA, P. & D.WEENIK (1992-2010): PRAAT. Doing phonetics by computer. Institute of Phonetic Sciences. Univ. Of Amsterdam. <http://www.praat.org>.

CABEDO, Adrián (2009): *La segmentación prosódica en español coloquial*, Valencia, Universitat de Valencia, anejo 67 de la revista Quaderns de Filologia.

CANTERO SERENA, Francisco José (2002): Teoría y análisis de la entonación. Barcelona, Edicions de la Universitat de Barcelona

FONT, Dolors (2007): *L'entonació del català*, Barcelona, Publicacions de l'Abadia de Montserrat, Biblioteca Milà i Fontanals, 53.

MAS MANCHÓN, Lluís (2008): “Testeo de 3 procedimientos de obtención del pitch para la modelización prosódica del discurso noticia”, *PHONICA*, vol. 4.

·
 NAVARRO TOMÁS, T. (1944): *Manual de entonación española*. New York: Hispanic Society (1974⁴) Madrid, Guadarrama.

QUILIS, Antonio (1981): *Fonética acústica de la lengua española*. Madrid, Gredos.

9. Anexo: corpus

Código	Enunciado
Ai-07	¿Y bambas tal vez no?
Ai-08	Llevaba gafas.
Ai-09	¿Llevaba gafas?
Ai-10	¿Y te fijaste si el pantalón tejano era elástico?
Ai-17	Tú eres Isabel.
Ai-18	Es tuyo
Ai-19	¿De José Antonio?
Ai-22	Que será lo mismo ¿no?
Ai-23	¿Pues ella también se lo imaginará supongo no?
Ai-26	No, no es desconocido para usted ¿verdad?
Ai-28	¿Los vándalos?
Ai-32	La o
AN-03-01-02	¿Te acuerdas de la caseta de Paco Lastras?
AN-03-03-01	¿A ti te gusta "Ojos verdes"?
AN-23-04-06	¿Tienes hijos?
AN-24-02-06	¿Estás grueso?
AN-25-03-01	¿También está puesto ahí?
AN-25-05-03	¿Me puedo levantar y cantar ya?
AN-26-01-06	¿Te hago una tila?
AN-26-02-08	Pero, ¿sabes bailar?
AN-27-02-08	¿Tiene usted hijos?
AN-40-03-09	Qué vives solo, Antonio.
AN-40-04-05	¿Y usted me fía?
AN-63-01-07	Y de eso ha mejorado, no, Antonia.
AN-71-01-07	¿Te parece bien?
AN-81-02-03	¿Quieres un saco de papas, cinco euros?
AN-82-06-01	Me sigues
AN-85-01-07	Vale, lo cojo ya
AN-85-02-02	Eh, por tal de desbloquear esta situación haría una coalición con el partido socialista
AR-01-01-02	Y van a llegar a un acuerdo
AR-01-01-04	Va a comprar alguno
AR-02-01-04	Sabe además que por lo menos hay cuatro mil agricultores que todavía siguen produciendo maíz no transgénico.
AS-04-06-03	No se cuela ninguna
CA-08-02-04	Mi niño y tú quieres que yo me lo coma aquí

CA-15-02-03	Y se gastó mil trescientos cincuenta y dos euros en ocho meses
CA-17-03-02	Me entiendes
CA-39-04-01	Hubo miedo
CA-40-04-03	Ustedes estarían por la hospitalización
CA-40-05-05	Tienen dinero para coger las armas
CA-43-03-03	La sueles tener baja
CA-59-02-01	Un día de trabajo nada más
CM-02-04-04	¿La asociación lo rehabilitó?
CM-10-05-01	¿Me lo puedes repetir?
CM-10-05-04	¿Corto yo también?
CM-11-03-02	¿Es bonito?
CM-17-03-03	Era muy jovencito, ¿verdad?
CM-18-02-06	¿Y por aquí tenéis una cañada real, verdad?
EX-01-01-04	¿Conocéis gente de Portugal que venga para acá a trabajar?
EX-01-01-05	¿Usted conoce gente de allí, de Portugal?
EX-03-01-12	¿Vamos a buscarlo?
EX-10-01-01	¿Veis el buitre?
EX-11-05-05	¿Compraste algo?
EX-11-06-08	¿Entramos?
EX-13-01-03	¿Me entiendes?
EX-14-01-01	¿A ti te gusta que te den masajes?
EX-21-01-01	¿Sabes lo que es un brochón, no?
EX-41-03-01	A día de hoy
EX-41-03-06	¿A tu favor?
EX-44-03-09	¿Eso se está haciendo?
EX-46-09-03	¿Te atreves?
EX-46-09-05	¿Tienes permiso de conducir?
MA-02-03-01	¿A Olegario lo conocéis, es un vecino de toda la vida?
MU-01-11-05	¿Te gusta el queso?
MU-01-11-06	¿Te gusta el tomate?
MU-03-03-03	Os habéis fijado en la Virgen de la esquina.
MU-10-02-05	¿Te cabe todo ahí?
MU-10-03-01	¿Es una indirecta para que me vaya?
MU-10-04-05	¿Alguien me llama por teléfono?
MU-13-04-06	¿Estás seguro?
MU-13-04-09	Entonces, ¿me puedo quitar la corbata también?
MU-15-01-01	¿Árbol incluido?
MU-59-01-03	¿Salgo guapa o no?

MU-76-01-10°	Si queréis verla funcionar
MU-76-08-02	Hola, me puede dar el premio Planeta que me han encargado
MU-81-02-03	Este, este tema, ¿tú lo has cantado por el tema de Julio Madrid?
MU-98-03-05	Tú llegaste a conocer al Campusino
PV-01-02-01	Ya estáis preparados
PV-01-12-01	Nos sacas una tanga para ver cómo es
PV-01-23-02	Te lo has probado
PV-03-05-04	Puedes decir algo