

VARIABLES PROSÓDICAS EN LA IDENTIFICACIÓN DEL LOCUTOR

Josefa Dorta y Chaxiraxi Díaz

Laboratorio de Fonética (SEGAI de la Universidad de La Laguna)

jdorta@ull.edu.es/ chadiaz@ull.edu.es

Resumen: La F0 es un parámetro de especial interés en la praxis forense puesto que, entre otras ventajas, se puede analizar incluso en grabaciones deficientes (Rose, 2002). Por otra parte, las tesis variacionistas labovianas (Labov, 1994, 2001) indican que la variación es inherente en todas las lenguas y se produce tanto por factores internos como externos. Partiendo de estos supuestos, analizaremos un conjunto de oraciones controladas emitidas por sujetos con características sociolingüísticas similares. Consideraremos las configuraciones tonales (F0 máxima y mínima, media, desviación estándar, asimetría y curtosis), las pendientes tonales de los acentos y el alineamiento de los picos. Los resultados servirán para afianzar hasta qué punto es útil el estudio de las variables consideradas en un peritaje forense con el propósito de aportar información sobre el locutor y su procedencia geográfica.

Palabras clave: Fonética forense; configuraciones tonales; picos tonales; pendientes tonales; análisis discriminante.

Abstract: The F0 is a parameter of special interest in the forensic praxis due to the fact that, among other advantages, it can be analyzed even on low quality recordings (Rose, 2002). On the other hand, Labov's variationist approach (Labov, 1994, 2001) indicates that the variation is inherent in all languages and is produced by internal and external factors. Starting from these assumptions, we will analyze a set of controlled sentences uttered by speakers with similar sociolinguistic characteristics. We will consider the tonal configurations (F0 maximum and minimum values, mean, standard deviation, skewness and kurtosis), the tonal slopes of pitch accents and the alignment of the peaks. The results will show the usefulness of the considered variables in forensic speaker identification as well as his geographic source.

Keywords: Forensic phonetics; tonal configurations; tonal peaks; tonal slopes; discriminant analysis.

Josefa Dorta y Chaxiraxi Díaz

1. Introducción

La importancia de la frecuencia fundamental (F0) -y de otros aspectos relacionados con ella, como el contorno tonal y la alineación tonal- ha sido destacada en la praxis forense por diversos autores (Stevens, 1971; Atal, 1972; Brown, 1981; Nolan, 1983; Baldwin i French, 1990; Hollien, 1990; Braun, 1995; van Dommelen, 1997; Rose, 2002; Cicres, 2007; Albalá, Battaner, Carranza, De la Mota, Gil, Llisterri, Machuca, Madrigal, Marquina, Marrero, Riera & Ríos, 2008) pues se considera que contribuye a individualizar la voz, aspecto prioritario en la tarea de verificación e identificación forense de hablantes, ya que delata rasgos característicos como la edad, sexo, lengua o variedad lingüística, nivel sociocultural, etc. (Nolan, 2002; Escudero, Cardeñoso, Sánchez, Navas & Hernández, 2003; Cicres & Turell 2005)¹. La importancia de la F0 en el ámbito forense radica, además, en el hecho de que se transmite bien en condiciones poco favorables, como pueden ser llamadas telefónicas o ambientes con ruido, por lo que es un parámetro que puede analizarse incluso en grabaciones deficientes (Nolan, 1983; Braun & Köster, 1995; Rose, 2002; Cicres, 2003).

En diversos estudios se ha visto que la variación de la F0 es mayor entre hablantes (variación interhablante) que en un mismo hablante (variación intrahablante). Por otra parte, el estudio de este parámetro acústico para tratar de individualizar la voz con objetivos forenses se ha centrado fundamentalmente en la media y en la desviación estándar. No obstante, en un estudio relativamente reciente para el español, Albalá *et alii.* (2008) concluyen que si bien el valor medio de la F0 es determinante, pues es responsable de las variaciones registradas entre los informantes analizados, no sucede lo mismo con la desviación estándar.

Rouas, Farinas, Pellegrino & André-Obrecht (2003) investigaron con anterioridad, además de la media y de la desviación estándar, otras dos medidas de F0, esto es, el coeficiente de asimetría y la curtosis con el objetivo de identificar de manera automática lenguas diversas (inglés, alemán, español, francés, japonés, mandarín, vietnamita, tamil, persa y coreano). El análisis se centró en cada segmento y se obtuvo mejores resultados en lectura que en habla

¹ La F0 es el parámetro acústico que indica la frecuencia de vibración de las cuerdas vocales; como estas forman parte de la anatomía humana, se trata de un parámetro que se relaciona con la fisiología de cada hablante.

espontánea. En un estudio posterior, Kinoshita, Ishihara & Rose (2009) opinan que la asimetría y la curtosis pueden aportar buena información en japonés por lo que, siguiendo a estos autores, en una investigación más reciente, Marquina (2011) considera las cuatro variables mencionadas con el objetivo de comprobar su relevancia en el estudio de la variación inter e intralocutor y de la variación inter e intralingüística de hablantes bilingües catalanes y castellanos para ratificar si los resultados permiten individualizar las voces de los hablantes analizados con independencia de la lengua en que se comuniquen. En lugar de centrarse en el análisis de la F0 de los segmentos vocálicos, Marquina (2011:54-55) estudia las cuatro variables en grupos fónicos y concluye que dichas variables

contribuyen a la determinación de la individualidad de los hablantes bilingües equilibrados de catalán y de castellano cuando emplean la misma lengua e incluso, aunque en un porcentaje menor, independientemente de la lengua utilizada, puesto que la variación interlocutor es significativamente mayor que la variación intralocutor.

El presente trabajo obedece a dos necesidades que consideramos fundamentales: una, que en el ámbito del español es necesario seguir realizando investigaciones que contribuyan a la individualización de los hablantes por su voz; y dos, teniendo en cuenta las tesis variacionistas labovianas (Labov, 1994, 2001) que indican que la variación es inherente en todas las lenguas y se produce tanto por factores internos (fonético-fonológicos, sintácticos, semánticos, pragmáticos, etc.), como externos (sociales, geográficos, individuales, etc.), pensamos que es igualmente importante la realización de trabajos que se relacionen con diferentes variedades puesto que, además de la información individual, proporcionarán sin duda información sociolingüística relevante para la práctica forense.

2. Hipótesis

Asumimos que la F0 es un parámetro robusto que nos permite discriminar la voz de los locutores y que, siendo uno de los fundamentales que definen la entonación, es posible encontrar variables que permitan diferenciar dos voces. Asimismo, se espera que no todas las variables analizadas tengan la misma eficacia o peso en la

identificación forense de hablantes por parte de la autoridad judicial, por lo que el estudio aportará información sobre las más rentables para tal fin.

3. Objetivos

Esta investigación responde a la necesidad de realizar trabajos específicos, en este caso relacionados con aspectos entonativos de las variedades del español, con el propósito de que los resultados puedan contribuir a la praxis forense. En la presente investigación se ha elegido la variedad canaria para determinar la relevancia de las variables objeto de estudio en la individualización de hablantes masculinos y femeninos de esta variedad en un caso pericial de voz. En el análisis consideraremos, por una parte, las configuraciones tonales teniendo en cuenta la F0 máxima y mínima, la media, la desviación estándar, la asimetría y la curtosis. Además, tendremos en cuenta la sincronización de los picos inicial y final, sea este nuclear o no, así como las pendientes de los acentos inicial y final.

4. Metodología

4.1. Corpus e informantes

Con el propósito de controlar que la variación en las variables consideradas en el análisis está motivada más por cuestiones idiolectales que por otro tipo de factores, hemos elegido un corpus experimental de 216 oraciones breves de once sílabas emitidas, mediante elicitación textual, por hablantes con características sociolingüísticas similares, esto es, un grupo de 4 informantes canarios, dos hombres (H) y dos mujeres (M) de procedencia urbana, nacidos en San Cristóbal de La Laguna (TF) y Las Palmas de Gran Canaria (GC), con un nivel sociocultural bajo y con edades similares comprendidas entre 25 y 55 años. De este modo, las oraciones son idénticas en todos los informantes ² y las características

² El corpus procede del proyecto PI FFI2010-16993 *La entonación interrogativa y declarativa del español de Canarias y su relación con la de Cuba y Venezuela*, subvencionado por el Ministerio de Economía y Competitividad. En este proyecto, el corpus sigue las directrices del *Atlas Multimédia Prosodique de l'Espace Roman* (AMPER)[<http://www.u-grenoble3.fr/dialecto/AMPER/amper.htm>] dirigido

sociolingüísticas de estos muy similares. La elección de dos informantes por cada sexo obedece a la idea de emular una situación muy común en la práctica forense, esto es, la confrontación de dos voces (indubitada y dubitada) que a priori no podemos decidir si pertenecen o no a un mismo locutor. El corpus se distribuye como se puede ver en la tabla 1.

	Corpus			
	Voz femenina		Voz masculina	
	M TF	M GC	H TF	H GC
Declarativas	27	27	27	27
Interrogativas	27	27	27	27
n= 216	54	54	54	54

Tabla 1

4.2. Análisis

En primer lugar, todas las oraciones, previamente segmentadas con el programa GoldWave Digital Audio Editor (versión 4.25), fueron visualizadas y escuchadas con el programa Praat para Windows (5.3.63 de 24 de enero de 2014) con el objetivo de asegurar que la F0 no presentaba puntos dispersos que alterarían, sin duda, las medias y los valores del análisis. Posteriormente, se usó un script de Praat para extraer textgrid de todas las oraciones del corpus; a partir de los archivos de sonido y de los textgrid creados obtuvimos de manera automática los valores de F0 (media, máxima y mínima, desviación estándar, asimetría y curtosis) con el uso de un nuevo script³.

Los valores relacionados con la sincronización de los picos y las pendientes de los acentos han sido obtenidos con el programa AMPER-2006 creado a partir de rutinas de Matlab⁴ (López Bobo, Muñiz Cachón, Díaz Gómez, Corral Blanco, Brezmes Alonso & Alvarelos Pedrero, 2007); el análisis extrae tres valores de F0 en cada una de las vocales silábicas. Se ha tomado el valor central, por ser más

internacionalmente por Michel Contini y Antonio Romano. En este caso, las oraciones de 11 sílabas constan de un SN y de un SPrep trisílabos en los que se combinan los tres acentos del español (agudo, llano y esdrújulo).

³ Los scripts utilizados han sido creados por Mietta Lennes y Sandra Schwab. Agradecemos la colaboración prestada por Jordi Cicres y Sandra Schwab.

⁴ Licencia de Matlab n° 256105 del Laboratorio de Fonética de la ULL.

Variables prosódicas en la identificación del locutor

estable, para determinar la sincronización de este con las sílabas acentuadas o no.

Finalmente, todos los valores obtenidos han sido sometidos a un análisis estadístico discriminante cuya finalidad es analizar si existen diferencias significativas entre grupos respecto a un conjunto de variables medidas sobre los mismos. Este análisis intenta crear un modelo matemático que discrimine correctamente, con cierto porcentaje de error, entre los grupos, lo que permite extraer las conclusiones pertinentes encaminadas a la identificación de las voces analizadas.

5. Resultados

5.1. La configuración tonal

La configuración tonal de las oraciones analizadas es muy similar en los cuatro informantes elegidos teniendo en cuenta, por un lado, las oraciones declarativas (con final descendente) y, por otro, las interrogativas (con final ascendente-descendente o circunflejo)⁵. En la figura 1 se puede ver un ejemplo de ambas modalidades.

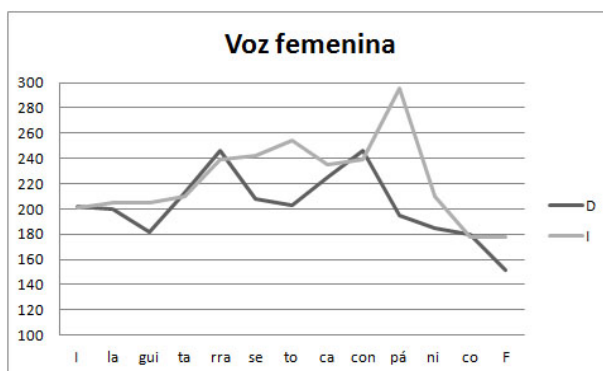


Figura 1

⁵ El final circunflejo ilustrado en la figura 1 es característico en las interrogativas absolutas canarias aunque en casos esporádicos se realice una entonación interrogativa con final ascendente que, según han revelado numerosos estudios previos que hemos realizado del español de Canarias, es fruto de una imitación del castellano en contextos formales o de lectura.

Los valores obtenidos han sido la media de F0, la desviación estándar (medida del grado de dispersión de los datos con respecto al valor promedio), la asimetría (identifica cómo se distribuyen los datos en relación con la media aritmética⁶), la curtosis o apuntamiento (mide el grado de concentración que presentan los valores en la región central de la distribución⁷) y la F0 máxima y mínima. La separación de las oraciones por modalidad y sexo arrojó los resultados generales de la tabla 2.

	Mujeres				Hombres			
	Declarativas		Interrogativas		Declarativas		Interrogativas	
	TF	GC	TF	GC	TF	GC	TF	GC
Media de F0	195,2	200,9	188,2	190,6	168,9	173,2	173,6	191,9
F0 sd	35,5	22,6	36,1	37,3	17,2	32,4	13,8	39,5
Asimetría	-0,7	-0,4	0,3	0,4	-0,4	0,2	-0,2	-0,7
Curtosis	-0,1	0,5	0,6	0,5	-0,6	1,1	0,1	-0,9
F0 máx	251,7	246,3	254,0	258,1	198,9	224,4	203,9	255,8
F0 mín	111,3	141,0	109,8	115,2	134,1	117,9	144,4	111,9

Tabla 2

Los valores medios de F0 de nuestros informantes canarios son más altos que los ofrecidos por Tapias y García (2000): según estos autores la media de la frecuencia fundamental de la población

⁶ Una distribución es *simétrica* cuando los datos se distribuyen de manera equitativa a ambos lados de la media; cuando no sucede esto, la distribución es *asimétrica*, y puede ser *positiva* cuando la mayoría de los datos se distribuyen por encima de la media o *negativa* cuando los datos se concentran mayoritariamente por debajo de la media.

⁷ Hay tres tipos de distribución según su grado de curtosis: *leptocúrtica*, que presenta un elevado grado de concentración alrededor de los valores centrales de la variable; *mesocúrtica*, con una concentración normal y *platicúrtica*, con un reducido grado de concentración alrededor de los valores centrales de la variable. Una curtosis positiva refleja una distribución puntiaguda debido a una gran concentración de los valores mientras que una negativa indica una distribución más bien plana debido a la baja concentración de los valores. Cuando la distribución de los datos cuenta con un coeficiente de asimetría y de curtosis de ± 0.5 estamos ante una curva normal.

Variables prosódicas en la identificación del locutor

femenina y masculina adulta es de 180 y 125 Hz, respectivamente⁸. En los informantes canarios, en cambio, las dos mujeres alcanzan una media general de 198 Hz en declarativas y de 189 Hz en interrogativas; los hombres por su parte, arrojan una media de 171 y 182 Hz en las dos modalidades, respectivamente.

Para ver la relevancia estadística de las 6 variables de la tabla 2 se aplicó un análisis multivariante discriminante por modalidad y sexo para calcular las relaciones existentes entre dichas variables y su capacidad de discriminación. El objetivo de este tipo de análisis es encontrar la combinación de las variables que mejor permite discriminar a los grupos. Hallada dicha combinación (función discriminante), se ha utilizado para clasificar los datos. Los resultados son los que mostramos en los apartados siguientes.

5.1.1. Análisis estadístico en voz femenina

Para conocer si las variables que hemos considerado tienen o no poder discriminante contrastamos las pruebas de igualdad de las medias de los grupos (Lambda de Wilks). Teniendo en cuenta que las variables tienen poder discriminante si $p < 0,05$, observamos que en las declarativas solo presentan valores significativos la desviación estándar ($p = 0,000$) y la mínima de F0 ($p = 0,002$). Por tanto, en principio solo estas variables apuntan a que podemos discriminar las dos voces femeninas. En las interrogativas, en cambio, ninguna de las seis variables arrojó resultados significativos ($p > 0,05$) por lo que no parece tener sentido aplicar el análisis discriminante en esta modalidad; aún así lo hemos hecho con el objetivo de comprobar si el conjunto de las variables tiene o no poder discriminante.

Como era de esperar, el valor de correlación canónica fue mayor en las declarativas que en las interrogativas (0,732 vs. 0,577, respectivamente) por lo que el conjunto de las variables elegidas es más correcto para discriminar las voces femeninas en la primera modalidad que en la segunda. En ambas, no obstante, se encontró una función discriminante significativa que indicó la posibilidad de distinguir entre las dos voces con alguna probabilidad de éxito

⁸ Los valores ofrecidos por estos autores fueron obtenidos del corpus VOCATEL. Se trata de una base de datos realizada a partir de llamadas telefónicas a hombres y mujeres de todas las provincias españolas. Los locutores masculinos fueron 2.193 y los femeninos 2.316.

teniendo en cuenta el conjunto de las variables analizadas: en las declarativas [lambda de Wilks 0,464; $\chi^2(6) = 37,584$ p=0,000]; en las interrogativas [lambda de Wilks 0,667; $\chi^2(6) = 19,837$ p=0,003]. Ahora bien, el peso estadístico de las variables en la función discriminante utilizada en el análisis no es el mismo ni idéntico en las dos modalidades⁹ (figura 2): en las declarativas decrece desde la desviación estándar a la curtosis; en las interrogativas el peso es mucho menor que en las declarativas y el orden, de mayor a menor peso, varía ahora desde la F0 máxima hasta la curtosis, que sigue siendo la variable menos consistente.

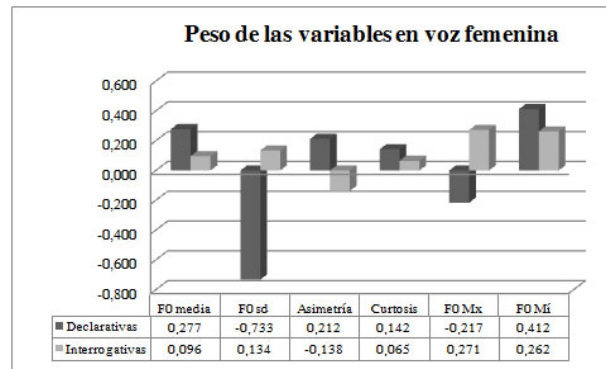


Figura 2

Finalmente, de acuerdo con el análisis discriminante realizado se interpretó correctamente el conjunto de casos analizados en porcentajes elevados. Así, del conjunto de declarativas agrupadas originalmente en la voz 1 (mujer de TF) y en la voz 2 (mujer de Gran Canaria) se clasificó correctamente un 81,5% y un 88,9%, respectivamente, con lo cual los errores de interpretación fueron pequeños por lo que, en general, en el 83,3% de los casos el análisis discriminante ha clasificado correctamente las dos voces. En las interrogativas, los porcentajes de interpretación correcta, aunque altos, fueron menos destacados que en las declarativas como presagiaba el

⁹ Los valores de peso estadístico se miden partiendo de 0, que representa el mínimo peso posible, hasta llegar a 1/-1, que indican el máximo peso en ambas direcciones. De esta forma un alto valor, tanto negativo como positivo, indica que la variable ha tenido mucho peso para hallar la función discriminante.

Variables prosódicas en la identificación del locutor

menor peso de las variables: 74,1% y 77,8% en la voz 1 y en la voz 2, respectivamente. En general, en esta modalidad el análisis discriminante ha clasificado correctamente los casos analizados en un 75,9%.

5.1.2. Análisis estadístico en voz masculina

En las declarativas de la voz masculina la Lambda de Wilks arroja resultados significativos ($p < 0,05$) en la desviación estándar, en la curtosis, en la F0 máxima y en la F0 mínima por lo que solo la F0 media y la asimetría no parecen tener valor discriminante. En las interrogativas, en cambio, a diferencia de la voz femenina, las seis variables presentan valores significativos. El valor de correlación canónica fue altísimo (0,927 y 0,975 en declarativas e interrogativas) y la función discriminante fue significativa en ambas modalidades: [lambda de Wilks 0,140; $\chi^2(6) = 96,240$ $p = 0,000$] y [lambda de Wilks 0,049; $\chi^2(6) = 148,173$ $p = 0,000$], en declarativas e interrogativas, respectivamente. El peso de las seis variables consideradas no fue idéntico en las dos modalidades, aunque en ambas desciende desde la desviación estándar a la asimetría (figura 3).

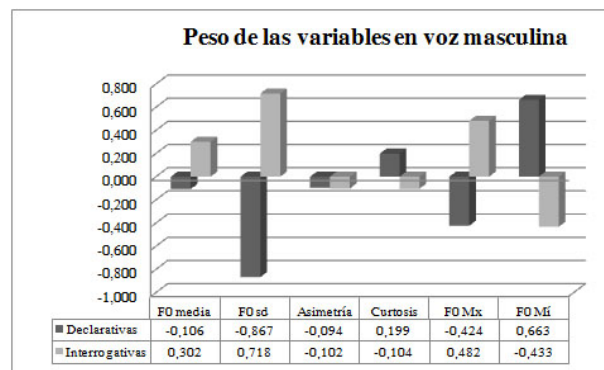


Figura 3

Destacamos, por último, que la discriminación fue correcta en el 100% de los casos agrupados originales y, por tanto, no se dio ningún error en ninguna de las dos modalidades.

5.2. Alineación y pendientes de los picos

En nuestro análisis hemos considerado, por una parte, la sincronización de los picos de F0 respecto de las sílabas acentuadas, esto es, del pico inicial de las declarativas e interrogativas, del último pico de las declarativas y del pico nuclear de las interrogativas; por otra, las pendientes de los acentos inicial y final.

5.2.1. Datos generales

5.2.1.1. Sincronización de los picos

Como es sabido, se pueden dar tres tipos generales de sincronización del pico tonal:

- a) Se alinea con la sílaba tónica (sincronización).
- b) Recae en una sílaba anterior a la tónica (prerealización).
- c) Coincide con una sílaba posterior a la tónica (posrealización).

Nuestros resultados generales indican que en las dos voces y modalidades estudiadas predomina en el pico inicial una sincronización con la postónica (posrealización) como se evidencia en los porcentajes de la figura 4, aunque es destacable que en el hombre de Gran Canaria (H GC) los porcentajes de sincronización con la tónica y la postónica aparecen prácticamente igualados.

Los casos que hemos clasificado como postónica en la figura 4 corresponden tanto a los que el pico se alinea propiamente con dicha

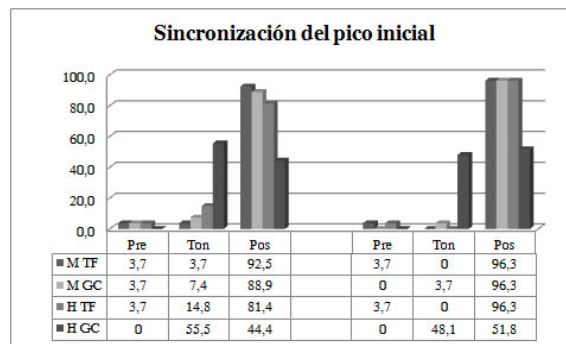


Figura 4

Variables prosódicas en la identificación del locutor

sílaba como aquellos en que lo hace con otra posterior. Así, en la figura 5 se pueden ver los porcentajes de alineamiento del pico inicial de las dos modalidades en las emisiones de los cuatro informantes. Como puede apreciarse, en las declarativas los porcentajes aparecen muy igualados en postónica y otra sílaba posterior; en cambio, en las interrogativas la tendencia mayoritaria es posponer el pico más allá de la postónica.

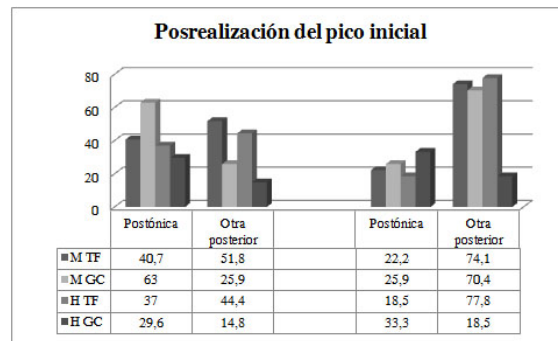


Figura 5

Por otra parte, las declarativas canarias se caracterizan, de manera general, por un descenso relativamente progresivo de la F0 desde un segundo pico (figura 1) que en el corpus analizado recae mayoritariamente en la postónica del verbo (sílabas 7) -marcando así una frontera sintagmática (SV/SPrep)- o en otra posterior (figura 6).

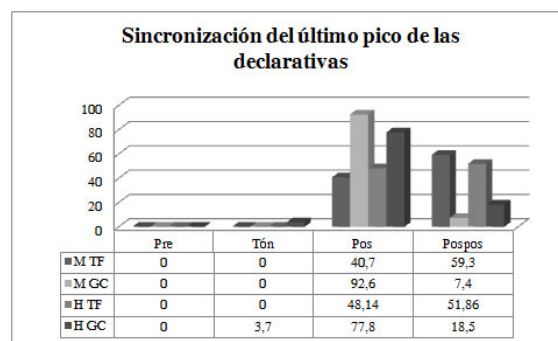


Figura 6

En lo que respecta a las interrogativas absolutas canarias, como ya se dijo (nota 5), tienen un final característico llamado comúnmente circunflejo, es decir, el que registra en el núcleo entonativo un movimiento de ascenso-descenso (figura 1). Como se puede ver en la figura 7, el pico nuclear se alinea mayoritariamente con la sílaba acentuada.

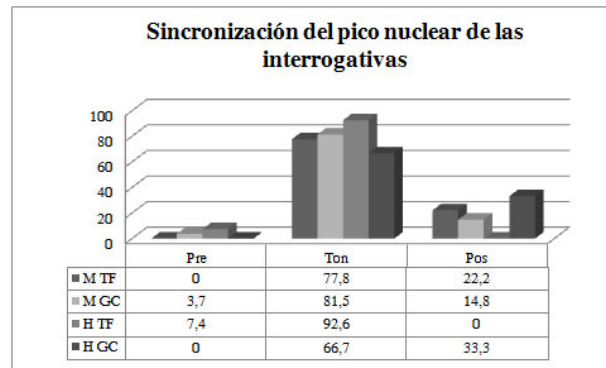


Figura 7

5.2.1.2. Pendientes inicial y final

Hemos considerado dos pendientes en las declarativas (figura 8), esto es, la del acento inicial generada desde el valor más bajo o valle inicial hasta el primer pico (pendiente 1) y la correspondiente al descenso desde el segundo pico hasta el final (pendiente 2).

En las interrogativas, dado el final circunflejo ya comentado, hemos considerado tres pendientes (figura 9): la del acento inicial, como en las declarativas; la del valle-pico nuclear (pendiente 2) y la que marca el descenso desde el pico hasta el final (pendiente 3).

A la luz de los datos de las figuras 8 y 9 se concluye que si consideramos el umbral psicoacústico de 1,5 St las pendientes analizadas en las dos modalidades presentan valores que superan dicho umbral. El análisis estadístico que veremos a continuación nos indicará si estas variables tienen suficiente peso para discriminar las voces entre sí.

Variables prosódicas en la identificación del locutor

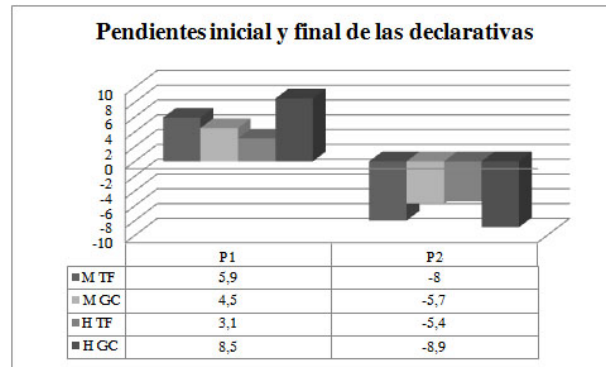


Figura 8

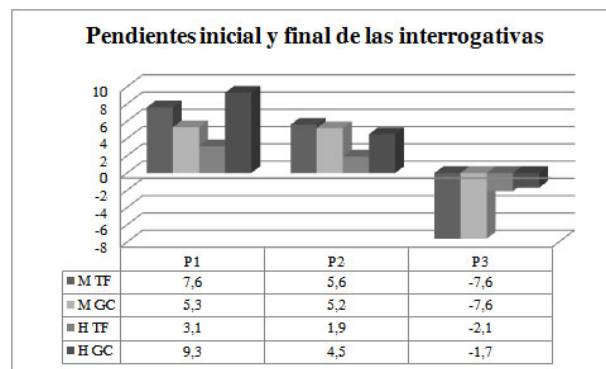


Figura 9

5.2.2. Análisis estadístico

Los valores de sincronización de los picos y de las pendientes fueron sometidos a un análisis estadístico multivariante de tipo discriminante como el aplicado a la configuración tonal (apartados 5.1.1. y 5.1.2.). Los resultados son los que se muestran a continuación.

5.2.2.1. Análisis estadístico en voz femenina

En las declarativas, las variables independientes del análisis discriminante fueron cuatro: sincronización del pico inicial (Sin P1),

sincronización del último pico (Sin P2), la pendiente inicial (P1) y la pendiente final (P2). En las interrogativas, como se ha dicho ya, fueron cinco: Sin P1, Sin P2, P1, P2¹⁰ y P3¹¹. Las pruebas de igualdad de las medias de grupos indicaron que en las declarativas la sincronización de los picos no presenta valores significativos en Sin P1 ($p=0,223$) pero sí en Sin P2 ($p=0,034$), así como en las dos pendientes ($p=0,000$ en ambos casos). En las interrogativas, en cambio, solo tiene valor significativo la pendiente inicial ($p=0,000$). Por tanto, las variables consideradas tienen mayor poder discriminante en las declarativas, donde el valor de correlación canónica fue más alto que en la otra modalidad (0,656 y 0,568, respectivamente). En ambas modalidades se encontró una función discriminante significativa que posibilita la diferenciación de las dos voces: en las declarativas [lambda de Wilks 0,570; $\chi^2(4) = 28,112$ $p=0,000$]; en las interrogativas [lambda de Wilks 0,677; $\chi^2(5) = 19,313$ $p=0,002$]. De las cuatro variables consideradas en las declarativas, la que tiene mayor peso estadístico es la pendiente final (P2) y la que menos la sincronización del pico inicial (Sin P1), como ya apuntaban los valores significativos de las pruebas de igualdad. En las interrogativas, en cambio, la variable de más peso es la P1 y la de menor peso es la Sin P1, que en este modelo no tiene peso alguno (figura 10).

Finalmente, según el análisis discriminante, de los casos agrupados originalmente como voz1 (TF) y como voz2 (GC), fueron interpretados correctamente en las declarativas un 70,4% y un 81,5%, respectivamente, con un porcentaje general de 75,9%. Por tanto, el modelo de análisis permite discriminar las dos voces en un porcentaje muy elevado. En las interrogativas los porcentajes, aunque altos, descendieron respecto de las declarativas: fueron clasificados correctamente un 66,7% y un 77,8% de los casos originalmente asignados a la voz 1 y 2, respectivamente con lo que, en general, fueron clasificados correctamente un 72,2% de los casos originales.

¹⁰ Recuérdese que la P2 en las interrogativas no es la misma que en las declarativas pues en estas es la generada por último pico-final absoluto mientras que en las interrogativas la P2 es la anterior (V-Pico) del acento nuclear.

¹¹ La P3 de las interrogativas coincidiría con la P2 de las declarativas en el hecho de que es la pendiente pico-final absoluto; no obstante, en las declarativas no se trata del pico nuclear mientras que en las interrogativas sí.

Variables prosódicas en la identificación del locutor

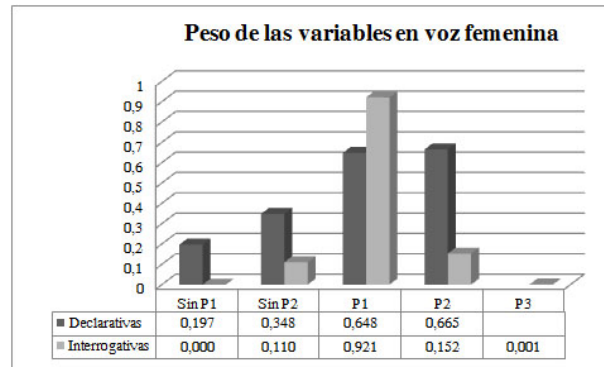


Figura 10

5.2.2.2. Análisis estadístico en voz masculina

De las cuatro variables consideradas en las declarativas, las pruebas de igualdad de las medias de los grupos muestran que la sincronización de los picos (Sin P1 y Sin P2) no tiene valor significativo. Por tanto, las únicas variables que apuntan a una buena discriminación de las voces son P1 y P2 ($p=0,000$). En las interrogativas, al contrario que en voz femenina, todas las variables mostraron un valor significativo, excepto P3 ($p=0,381$). El valor de correlación canónica fue altísimo, aunque superior en las declarativas (0,917 frente a 0,890 en las interrogativas) con lo cual se apunta en ambas modalidades a una buena discriminación de las voces; la Lambda de Wilks fue significativa [0,159; $\chi^2(4) = 91,922$ $p=0,000$], en declarativas y [0,208; $\chi^2(5) = 77,727$ $p=0,000$] en interrogativas. De las cuatro variables, la que mayor peso estadístico tiene en la primera modalidad es la pendiente 1 (P1) y la de menor peso es la sincronización del último pico (Sin P2); en las interrogativas, en cambio, la de mayor peso es la P2 y la de menor peso es la P3 (figura 11).

Los resultados de clasificación fueron muy altos e idénticos en las dos modalidades. Así, de los casos iniciales, fueron clasificados correctamente el 100% en la voz de TF y un 96,3% en la de GC; en general el porcentaje de correcta clasificación fue de un 98,1%.

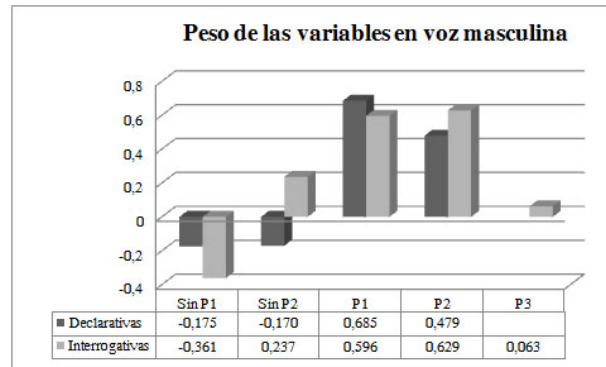


Figura 11

6. Conclusiones parciales

1.^a Se ha evidenciado la utilidad de separar el análisis de la voz femenina de la masculina y la modalidad declarativa de la interrogativa puesto que las variables analizadas, como preveía nuestra hipótesis, no tienen la misma relevancia y, por tanto, los mismos resultados en ambas voces y modalidades.

2.^a En lo que respecta a la configuración tonal, el análisis estadístico ha mostrado que el modelo de variables elegido fue menos adecuado en la voz femenina que en la masculina y, por lo que respecta a la modalidad, menos idóneo en las interrogativas de la voz femenina que en las declarativas; en cambio, en voz masculina dicho modelo es bastante adecuado en ambas modalidades. Hay que destacar que, tanto en voz femenina como masculina, la desviación estándar es la variable con mayor peso discriminante, con la única excepción de las interrogativas de las mujeres donde la que tiene mayor peso es la F0 máxima. Las variables de menor peso son la curtosis (en voz femenina) y la asimetría (en voz masculina). Destacamos igualmente que en ningún caso la F0 media ha sido la variable de mayor peso discriminante por lo que nuestros resultados son contrarios a los obtenidos por Albalá *et alii*. (2008). El análisis discriminante clasificó mejor los casos previamente asignados en voz masculina (100%) que en voz femenina y en esta resultó mejor la clasificación en las declarativas (83,3%) que en las interrogativas (75,9%).

3.^a En cuanto a la sincronización de los picos, se comprobó que el pico inicial de las declarativas e interrogativas y el segundo pico de las primeras se posponen mayoritariamente (posrealización), mientras que el pico nuclear de las interrogativas se alinea normalmente con la tónica nuclear. Este comportamiento, regular en hombres y mujeres, determina que las pruebas de igualdad de las medias de grupos solo arrojen resultados significativos en la variable Sin P2 (sincronización del pico 2) de las declarativas en voz femenina y en Sin P1 y Sin P2 de las interrogativas en voz masculina. Por tanto, estas variables no parecen tener suficiente poder discriminante.

4.^a Las pendientes generan en todos los casos valores que superan el umbral psicoacústico de 1,5 St y estadísticamente arrojaron valores significativos en las pruebas de igualdad de las medias de grupos por lo que, al contrario de lo que sucede con la sincronización de los picos, parecen tener valor discriminante. Las únicas excepciones se dieron en la P2 y P3 de las interrogativas puesto que ninguna de las dos pendientes alcanzó valores significativos en voz femenina, mientras que en los hombres no llegó a tener valor significativo la última. La variable de más peso varía según la modalidad y el sexo: en las declarativas de las mujeres, es la segunda pendiente (P2), mientras que en las interrogativas es la P1. En voz masculina sucede exactamente lo contrario. Por último, en las mujeres, los casos asignados originalmente a la voz1 (TF) y a la voz2 (GC) fueron clasificados correctamente en un 75,9% y en un 72,2% en declarativas e interrogativas, respectivamente. Los porcentajes de acierto fueron mayores en voz masculina: 100% y 98,1%% en declarativas e interrogativas, respectivamente.

7. Conclusión general

El estudio realizado evidencia la utilidad de realizar investigaciones sobre variedades específicas de una lengua con objetivos forenses. Así, por ejemplo, se ha podido comprobar que los valores medios de F0 de los informantes canarios son más altos que los ofrecidos por Tapias y García (2000). Por otra parte, se hace necesario comprobar si al eliminar las variables que en este trabajo no han evidenciado un

gran poder discriminante en voz masculina y/o femenina mejora el modelo de análisis para clasificar nuevos casos.

8. Referencias bibliográficas

- Albalá, María José; Battaner, Elena; Carranza, Mario; De la Mota, Carmen; Gil, Juana; Llisterri, Joaquín; Machuca, María Jesús; Madrigal, Natalia; Marquina, Montserrat; Marrero, Victoria; Riera, Montserrat & Ríos, Antonio. 2008. "VILE: Análisis estadístico de los parámetros relacionados con el grupo de entonación". *Language and Design. Journal of Theoretical and Experimental Linguistics, Special Issue 2: Experimental Prosody* (2): 15-22.
- Atal, Bishner. 1972. "Automatic speaker recognition based on pitch contours". *JASA* 52: 1687-1697.
- Baldwin, John & French, Peter. 1990. *Forensic Phonetics*. New York: Pinter Publishers.
- Braun, Angelika. 1995. "Fundamental frequency – how speaker-specific is it?". En Braun, Angelika & Köster, Jens-Peter (eds.) *Studies in Forensic Phonetics*. Trier: Wissenschaftlicher Verlag (Beiträge zur Phonetik und Linguistik), 9-23.
- Brown, Roger. 1981. "An experimental study of the relative importance of acoustic parameters for auditory speaker recognition". *Language and Speech* 24 (4): 295-310.
- Cicres, Jordi. 2003. "Estudi pilot d'identificació de parlants amb finalitat forense". *Papers de l'IULA. Sèrie Monografies* 7. Barcelona: Universitat Pompeu Fabra.
- Cicres, Jordi. 2007. *Aplicación del análisis de la entonación y de la alineación tonal a la identificación de hablantes en fonética forense*. Tesis doctoral. IULA: Universidad Pompeu Fabra [http://prosodia.upf.edu/home/arxiu/publicacions/recursosatles/cicres_aplicacio-analisi-entonacio.pdf].
- Cicres, Jordi & Turell, María Teresa. 2005. "Short and long-term variation in intonation patterns: a preliminary study for speaker identification". En *7th Biennial Conference on Forensic Linguistics/Language and Law* (Cardiff University, Cardiff [Regne Unit], 1-4 de julio).
- Escudero, David; Cardeñoso, Valentín; Sánchez, Jon; Navas, Eva & Hernández, Imma. 2003. "Uso de entonación en reconocimiento Automático de Locutor: resultados preliminares". En *Actas del II Congreso de la Sociedad Española de Acústica Forense*. Barcelona: Ceysa, 167-174.
- Hollien, Harry. 1990. *The Acoustics of Crime. The New Science of Forensic Phonetics*. New York: Kluwer Academic-Plenum Publishers (Applied Psycholinguistics and Communication Disorders).

Variables prosódicas en la identificación del locutor

- Kinoshita, Yuko; Ishihara, Shunichi & Rose, Philip. 2009. "Exploring the discriminatory potential of F0 distribution parameters in traditional forensic speaker recognition". *The International Journal of Speech, Language and the Law* 16 (1): 91-111.
- Labov, William. 1994. *Principles of Linguistic Change, I: Internal Factors*. Oxford: Blackwell.
- Labov, William. 2001. *Principles of Linguistic Change, II: External Factors*. Oxford: Blackwell.
- López Bobo, María Jesús; Muñiz Cachón, Carmen; Díaz Gómez, Liliana; Corral Blanco, Norberto; Brezmes Alonso, David & Alvarelos Pedrero, Mercedes. 2007. "Análisis y representación de la entonación. Replanteamiento metodológico en el marco del proyecto AMPER". En Dorta, Josefa (ed.) *La prosodia en el ámbito lingüístico románico*. Madrid-Santa Cruz de Tenerife: La Página Ediciones, Colección Universidad, 17-34.
- Marquina Zarauza, Montserrat. 2011. "Estudio acústico de la variación inter e intralocutor en la frecuencia fundamental de hablantes bilingües de catalán y de castellano". Trabajo de investigación de doctorado. [http://prosodia.upf.edu/home/arxiu/tesis/master/Marquina_Zarauza.pdf]
- Nolan, Francis. 1983. *The Phonetic Bases of Speaker Recognition*. Cambridge: Cambridge University Press.
- Nolan, Francis. 2002. "The 'telephone effect' on formants: a response". *Forensic Linguistics. The International Journal of Speech, Language and the Law* 9 (1): 74-82.
- Rose, Philip. 2002. *Forensic Speaker Identification*. London: Taylor & Francis (Taylor & Francis Forensic Science Series).
- Rouas, Jean-Luc; Farinas, Jérôme; Pellegrino, François & André-Obrecht, Régine. 2003. "Modelling prosody for language identification on read and spontaneous speech". En *International Conference on Acoustics, Speech and Signal Processing* 1, 40-43.
- Stevens, Kenneth. 1971. "Sources of Inter and intra-speaker variability in the acoustic properties of speech sounds". En Charboneau, Rene & Rigault, Andre (eds.) *Proceedings of the Seventh International Congress of Phonetic Sciences*. The Hague: Mouton, 206-232.
- Tapias Merino, Daniel & García, Carlos. 2000. "La frecuencia fundamental de la voz y sus efectos en el reconocimiento de habla continua". *Procesamiento del lenguaje natural* 26: 163-168.
- Van Dommelen, Wim. 1997. "The contribution of speech rhythm and pitch to speaker recognition". *Language and Speech* 30 (4): 325-338.
- Wolf, Jared. 1972. "Efficient acoustic parameters for speaker recognition". *JASA* 51: 2044-2056.