



Revista Signos

ISSN: 0035-0451

revista.signos@ucv.cl

Pontificia Universidad Católica de Valparaíso
Chile

Cicres, Jordi

Comparación forense de voces mediante el análisis multidimensional de las pausas llenas

Revista Signos, vol. 47, núm. 86, diciembre-, 2014, pp. 365-384

Pontificia Universidad Católica de Valparaíso

Valparaíso, Chile

Disponible en: <http://www.redalyc.org/articulo.oa?id=157032730002>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto



Comparación forense de voces mediante el análisis multidimensional de las pausas llenas

Forensic voice comparison by means of the multidimensional analysis of filled pauses

Jordi Cicres

UNIVERSITAT DE GIRONA
ESPAÑA
jordi.cicres@udg.edu

Recibido: 31-V-2013 / **Aceptado:** 13-XII-2013

Resumen

En fonética forense, uno de los parámetros utilizados habitualmente en las comparaciones forenses de voces son las pausas llenas (sonidos dubitativos). El objetivo de este estudio es evaluar su potencial discriminante y, en consecuencia, su utilidad en la comparación forense de voces con el fin de lograr identificaciones de voz robustas. Para ello, se han analizado variables relacionadas con la estructura de formantes y la cualidad de voz de esta clase de pausas. Los resultados muestran que las variables referentes a la cualidad de voz son más discriminantes que las variables referentes a su estructura de formantes. Además, el estudio conjunto de todas las variables logra mejores resultados que cuando las variables se analizan independientemente. En el caso de utilizar un corpus indubitado de tipo cerrado se logra un 90% de clasificaciones correctas.

Palabras Clave: Fonética forense, pausas llenas, identificación de hablantes, sonidos dubitativos, comparación forense de voces.

Abstract

In forensic phonetics, one parameter commonly used in forensic voice comparisons is filled pauses (hesitant sounds). The aim of this study is to evaluate their discriminatory potential and, consequently, their usefulness in forensic voice comparisons so as to obtain reliable and robust voice identifications. In order to carry out this assessment, several variables related to the voice quality and the formant structure of these pauses are analyzed. The results show that the variables linked to the voice quality are more discriminatory than the variables related to the formant structure. Furthermore, the joint consideration of all the variables has achieved even better results than when those variables are considered separately. When a closed type known corpus is used, an average of 90% of correct classifications are obtained.

Key Words: Forensic phonetics, filled pauses, speaker identification, hesitant sounds, forensic voice comparison.

INTRODUCCIÓN

La lingüística forense se ha definido tradicionalmente como la interfaz entre lengua y derecho, que aúna la amplia tradición del estudio del lenguaje jurídico y judicial, por un lado, con el más reciente uso del análisis del lenguaje en el campo pericial. Entre otras obras de referencia, en Gibbons y Turell (2008) se propone una visión amplia de la lingüística forense (visión compartida también por la Asociación Internacional de Lingüistas Forenses) que se sustenta sobre tres ejes: el análisis del lenguaje jurídico y judicial; el estudio del lenguaje del procedimiento judicial; y el lenguaje como evidencia lingüística (en el ámbito pericial). Este trabajo se centra en este último eje: el uso del lenguaje por parte de peritos lingüistas con el fin de aportar evidencia lingüística en un procedimiento judicial. Más concretamente, el presente estudio se ciñe al análisis acústico de las pausas llenas con el fin de evaluar su potencial discriminante en la comparación forense de voces, lo que eventualmente puede llevar a la identificación de una persona por su voz.

La tarea de comparar voces dubitadas (aquellas cuya atribución a hablantes concretos se pone en duda en el procedimiento judicial) con voces indubitadas (aquellas atribuidas a hablantes conocidos y aceptados por las partes) es sin duda compleja y multidisciplinar, puesto que tanto la realidad física de la voz como las elecciones lingüísticas son objetos de estudio variables.

Por un lado, todos los hablantes tienen un modo distinto de hablar y de escribir, lo que permite situarlos dentro de una comunidad lingüística determinada y describir su ‘idiolecto’ o ‘estilo idiolectal’. Este concepto:

“[...] would have to do primarily, not with what system of language/dialect an individual has, but with a) how this system, shared by lots of people, is used in a distinctive way by a particular individual; b) the speaker/writer’s production, which appears to be ‘individual’ and ‘unique’ (Coulthard, 2004) and also c) Halliday’s (1989) proposal of ‘options’ and ‘selections’ from these options” (Turell, 2010: 217).

En otras palabras, el idiolecto explica la “variation within a language that is associated with individual speakers” (Burridge & Mulder, 1998: 302). Se presume que los hablantes tienen una forma única, idiosincrásica de utilizar el lenguaje. Es decir, que no se puede relacionar con factores externos, tales como el sexo, la edad, el dialecto, el sociolecto, etc. (Brown, 1982).

Por otro lado, la variabilidad de la voz viene determinada por una combinación de las características fisiológicas del hablante y fenómenos lingüísticos y paralingüísticos. Así, por una parte los hablantes muestran diferencias en las propiedades acústicas de sus producciones lingüísticas: cualidad de voz, frecuencia fundamental, estructura espectral y patrones de duración e intensidad de las consonantes y vocales (entre otros, Rose, 2002; Hollien, 2002). Por otra parte, los hablantes toman decisiones conscientes o inconscientes acerca del uso y elecciones en todos los niveles lingüísticos (Coulthard, 1994; Turell, 2010). Además, también eligen en ámbitos no estrictamente lingüísticos, como las pautas de respiración (Kienast & Glitza, 2003), la velocidad de articulación y ritmo, y el uso de pausas y sonidos no léxicos (como los que indican falta de fluidez en el habla, una confirmación, un canal de retorno, una respuesta, etc., además de las pausas llenas). Los sonidos no léxicos son muy frecuentes en el habla espontánea; según el estudio de Ward (2006), aparecen una media de una vez cada 5 segundos.

En fonética forense, un objeto de estudio utilizado habitualmente en las comparaciones forenses de voces son las pausas. Se pueden distinguir varios tipos de pausas: silenciosas, de turno, de respiración, fonológicas y llenas (Rose, 2002). Dentro de este último grupo, los hablantes pueden seguir distintas estrategias para llenar el silencio: mediante la utilización de elementos léxicos como muletillas, alargamiento de los sonidos vocálicos o consonánticos que se hallan a final de palabra, o bien utilizando un sonido dubitativo (no léxico).

El interés forense en el análisis de las pausas llenas se debe a varios factores que han sido estudiados.

En primer lugar, a la preferencia individual de utilizar el mismo tipo de pausa llena:

“Individuals tend to be quite consistent in using ‘their’ respective personal variant of the hesitation sound, in particular with respect to the optional addition of a bilabial nasal consonant and the colour of the vocalic component”. (Künzel, 1997: 51)

Este mismo estudio cuantificó el porcentaje de pausas llenas en relación a la totalidad de pausas, y los resultados indicaron que había diferencias estadísticamente significativas entre hablantes (una mayor variación inter-hablante), mientras que las diferencias entre producciones de los mismos hablantes eran pequeñas (un bajo nivel de variación intra-hablante).

En segundo lugar, las pausas llenas, al encontrarse habitualmente entre pausas silenciosas, se muestran menos afectadas por los efectos de la coarticulación, de modo que son menos variables y más estables en sus propiedades acústicas. En este sentido, Foulkes, Carrol y Hughes (2004) analizaron la variabilidad de los tres primeros formantes de las pausas llenas y la compararon con la de las vocales léxicas. Sus resultados indican que en la mayoría de casos no existen diferencias estadísticamente significativas (66% para F1 y F2, y 54% para F3), mientras que los casos de hablantes que presentan más variabilidad en las pausas llenas que en las vocales léxicas son mucho menores (10% para F1 y F2, y 6% para F3) que los casos en que las vocales léxicas son más variables que las pausas llenas (24% para F1 y F2 y 40% para F3). Ward (2006) y Pätzold y Simpson (1995) también analizaron estas pausas llenas en inglés americano y alemán respectivamente desde el punto de vista acústico, e igualmente hallaron que los timbres vocálicos de estas pausas llenas son distintos y menos variables que los timbres de las vocales en elementos léxicos.

En tercer lugar, se trata de elementos paralingüísticos y, por tanto, están más estrechamente relacionados con el idiolecto de cada hablante. Sin embargo, varios estudios han hallado diferencias significativas relacionadas con variables sociolingüísticas (Erickson, 1979; Tannen, 1990; Mulac, Erlandson, Farrar, Hallett, Molloy & Prescott, 1998; Foulkes et al., 2004). Por ejemplo, en este último estudio se demuestra que la presencia de un segmento nasal en las pausas llenas es significativamente mayor que su ausencia en mujeres, en personas de clase media, en hablantes jóvenes y en las interacciones de estas variables ($p < 0,001$ en todos los casos) en inglés británico.

Por último, desde la óptica micro-acústica, en Cicres y Turell (2005) y en Cicres (2007) se compararon una serie de parámetros referentes a las propiedades relacionadas con la cualidad de voz (frecuencia fundamental, perturbación de frecuencia o *jitter*, interrupción de la voz, perturbación de amplitud o *shimmer*, ruido, temblor, subarmónicos y aperiodicidad) de las pausas llenas. Ambos estudios analizaron el grado de variación intra-hablante e inter-hablante, y los resultados mostraron que existían diferencias significativas entre hablantes en las variables relacionadas con la frecuencia fundamental, *shimmer*, *jitter*, ruido y aperiodicidad. En el resto de variables, no se hallaron diferencias significativas.

1. Metodología

1.1. Objetivos e hipótesis

En este estudio se analizan los formantes y una serie de parámetros relacionados con la cualidad de voz de las pausas llenas con el fin de evaluar su capacidad discriminante y, consecuentemente, su utilidad en la comparación forense de voces. Los objetivos específicos son los siguientes:

- a) Comparar la capacidad discriminante de los distintos parámetros acústicos relacionados con las pausas llenas: formantes y parámetros relacionados con la cualidad de voz.
- b) Evaluar su potencial de aplicación en la comparación forense de voces utilizando un corpus indubitado de tipo cerrado (Nolan, 1983), es decir, cuando se conoce que el hablante de la grabación dubitada es uno de los hablantes de las muestras indubitadas.

En la comparación forense de voces, la mayoría de expertos y laboratorios utilizan una combinación de análisis lingüístico y acústico (Gold & French, 2011), complementado en ocasiones con herramientas automáticas (Delgado, 2005; González & Lucena, 2005, para citar a investigadores españoles) basadas en el marco bayesiano. La elección de una metodología u otra depende no solo de los medios disponibles, sino también de la pregunta de investigación objeto del análisis. Así, dada una grabación dubitada (objeto de la pericia) y un conjunto de muestras de los sospechosos, el análisis discriminante se perfila como la herramienta más útil. Si se trata de comparar una muestra dubitada con las muestras indubitadas del sospechoso y decidir si pertenecen a la misma persona, la inclusión de los ratios de verosimilitud y los corpus de referencia complementan el análisis (no sin polémica acerca de su posibilidad de uso con garantías, como se muestra en French & Harrison, 2007; Rose & Morrison, 2009; French, Nolan, Foulkes, Harrison & McDougall, 2010). En este estudio, se ha diseñado el análisis partiendo de la primera situación. Así, relacionadas con los objetivos planteados, se plantean las siguientes hipótesis:

- a) La inclusión en el análisis discriminante de las variables relacionadas con la cualidad de voz mejorará los resultados respecto a los obtenidos mediante el análisis de formantes.
- b) Utilizando un corpus de referencia cerrado, se obtendrán porcentajes altos de clasificaciones correctas mediante el análisis discriminante.

1.2. Corpus

Se ha creado un corpus con 15 grabaciones correspondientes a 10 hablantes distintos. Los hablantes son 5 hombres y 5 mujeres adultos, de edades comprendidas

entre los 25 y los 45 años. Todos son catalanohablantes. 5 de las grabaciones se realizaron mediante grabadoras portátiles en contextos forenses reales y se analizaron como muestras dubitadas; otras 5, de los mismos hablantes, se obtuvieron a posteriori como muestras indubitadas; finalmente, se añadieron otras 5 grabaciones sin ninguna relación con las anteriores.

En todos los casos, las propiedades técnicas de las grabaciones eran similares. Aun así, todas las grabaciones se han unificado en sus principales características técnicas: códec de audio, frecuencia de muestreo, número de canales y profundidad de bits, con el fin de igualar en calidad global del sonido a la originalmente de peor calidad. Así, las características técnicas de los archivos analizados han sido las siguientes: formato mp3 (a 128 kbps), frecuencia de muestreo de 16 kHz, mono y 16 bits de profundidad. La Tabla 1 muestra un resumen de las grabaciones y sus códigos identificativos:

Tabla 1. Resumen de las grabaciones analizadas. Las celdas sombreadas corresponden a grabaciones forenses, y las que están en la misma fila son las versiones dubitada e indubitada de la misma voz.

Grabaciones indubitadas forenses y no forenses		Grabaciones dubitadas (forenses)	
Código	Sexo	Código	Sexo
indubitada 1	Hombre	dubitada 1	Hombre
indubitada 2	Mujer	dubitada 2	Mujer
indubitada 3	Hombre	dubitada 3	Hombre
indubitada 4	Mujer	dubitada 4	Mujer
indubitada 5	Hombre	dubitada 5	Hombre
indubitada 6	Mujer		
indubitada 7	Hombre		
indubitada 8	Mujer		
indubitada 9	Hombre		
indubitada 10	Mujer		

Para cada grabación se han identificado 10 pausas llenas (en total, el estudio ha tenido en cuenta 150 sonidos). Los segmentos analizados son sonidos dubitativos cuyo timbre se encuentra en el espacio comprendido entre los sonidos [e:] y [ə:], entre pausas y sin elemento nasal. Las grabaciones escogidas tenían un mínimo de 10 pausas llenas que cumplían con estas condiciones. Para evitar inexactitudes en la identificación y segmentación de las pausas llenas, se ha efectuado el análisis de los formantes y de la cualidad de voz en el fragmento correspondiente al 50% central de cada sonido dubitativo.

1.3. Variables analizadas

Las variables acústicas analizadas tienen que ver con a) la estructura de formantes de los sonidos dubitativos, y b) la cualidad de voz (González, Cervera & Miralles, 2002). Concretamente, para la estructura de formantes se han analizado los tres

primeros formantes (F1 a F3) y la distancia entre formantes (F2-F1, F3-F1 y F3-F2). Para la cualidad de voz se ha analizado un conjunto de variables relacionadas con la frecuencia fundamental (media, mediana, desviación estándar y valores mínimo y máximo), sonoridad (número y porcentaje de periodos sordos), *jitter* (local, rap, ppq y ddb), *shimmer* (local, dB, apq3, apq5, apq11, dda) y ruido (razón ruido-armónicos). Las variables acústicas, que se resumen en la Tabla 2, se han analizado mediante el programa de análisis acústico Praat, en su versión 5.3.39 (Boersma, 2001).

Tabla 2. Resumen de las variables estudiadas (coincidentes con las de Cicres, 2007).

Grupo de parámetros	Variable	Explicación (entre paréntesis, unidad de medida)
Referentes a la frecuencia fundamental (F0)	Media	Valor medio de F0 (Hz)
	Mediana	Mediana de F0 (Hz)
	Desviación estándar	Desviación estándar de F0 (Hz)
	Valor mínimo	Valor mínimo de F0 (Hz)
	Valor máximo	Valor máximo de F0 (Hz)
Referentes a la interrupción de la sonoridad y a la aperioidicidad	Porcentaje de periodos sordos	Porcentaje de periodos sordos (%)
	Grado de segmentos sordos	Porcentaje de marcos, de 1 milisegundo, analizados como sordos (%)
Referentes a la perturbación de frecuencia (<i>jitter</i>)	<i>Jitter</i> (local)	Variabilidad periodo-a-periodo, dividida por el valor medio del periodo (%)
	<i>Jitter</i> (local, absoluto)	Variabilidad periodo-a-periodo (segundos)
	<i>Jitter</i> (rap)	Variabilidad periodo-a-periodo con un suavizado de 3 periodos, dividida por el valor medio del periodo (%)
	<i>Jitter</i> (ppq)	Variabilidad periodo-a-periodo con un suavizado de 5 periodos, dividida por el valor medio del periodo (%)
	<i>Jitter</i> (ddp)	Diferencia absoluta media entre diferencias consecutivas entre periodos consecutivos, dividida por el periodo medio (%)
Referentes a la perturbación de amplitud (<i>shimmer</i>)	<i>Shimmer</i> (local)	Variabilidad de amplitud de periodos consecutivos dividida por la amplitud media (%)
	<i>Shimmer</i> (local, dB)	Variabilidad de amplitud en periodos consecutivos en valores absolutos (dB)
	<i>Shimmer</i> (apq3)	Variabilidad de amplitud pico a pico con un suavizado de 3 periodos, dividido por la amplitud media (%)
	<i>Shimmer</i> (apq5)	Variabilidad de amplitud pico a pico con un suavizado de 5 periodos, dividido por la amplitud media (%)
	<i>Shimmer</i> (apq11)	Variabilidad de amplitud pico a pico con un suavizado de 11 periodos, dividido por la amplitud media (%)
	<i>Shimmer</i> (dda)	Diferencia absoluta media entre diferencias consecutivas entre las amplitudes de periodos consecutivos (%)
Referentes al ruido	Razón ruido/armónicos	Cantidad de energía no armónica presente en el sonido respecto a la energía armónica producida por la vibración de las cuerdas vocales. Incluye toda la energía inarmónica presente en la onda sonora, independientemente de su origen (turbulencias, irregularidades en los periodos vibratorios, etc.) (dB)

En primer lugar, se ha comprobado la existencia de variables estadísticamente correlacionadas (correlación de Pearson y correlación de Kendall-Tau-b), con lo que su información acústica resulta redundante. Los niveles de significación de las correlaciones se muestran en la Tabla 3:

Tabla 3. Variables correlacionadas significativamente.

F3 – F1 y F3 – F2	*
Todas las variables relacionadas con la frecuencia fundamental (excepto el valor mínimo)	**
Número y porcentaje de periodos sordos e interrupciones de sonoridad	**
Todas las variables relacionadas con el <i>jitter</i>	**
Todas las variables relacionadas con el <i>shimmer</i>	**

* $p < 0.05$; ** $p < 0.01$

Para lograr los objetivos del estudio se ha realizado un análisis lineal discriminante (ALD) con el hablante como factor y las variables referentes a las propiedades acústicas de las pausas llenas como variables dependientes. El ALD es una técnica de análisis multivariante que permite descubrir las relaciones existentes entre un gran número de variables. El proceso consiste en analizar los datos para encontrar la función o combinación lineal de las variables (la función discriminante) que mejor permite explicar las diferencias entre los distintos grupos (en este artículo, los hablantes).

A raíz de los resultados de los test de correlación, se ha simplificado el ALD utilizando únicamente una variable para cada grupo de variables correlacionadas significativamente. La lista definitiva de variables incorporadas al ALD es la siguiente:

- F1, F2 y F3
- Distancias interformánticas
- Mediana de F0
- Valor mínimo de F0
- Porcentaje de segmentos sonoros
- *Jitter* absoluto
- *Shimmer* absoluto
- Razón ruido-armónicos

En este estudio, se ha optado por el método de incluir todas las variables independientes juntas. Sin embargo, no es hasta que se clasifican nuevos casos que se puede evaluar realmente la capacidad discriminante de las funciones. Para ello,

se pueden añadir nuevos casos conocidos, y comprobar si quedan bien clasificados según las funciones discriminantes, o bien se puede utilizar el método de clasificación dejando uno fuera (validación cruzada), que consiste en eliminar casos uno a uno del modelo y clasificarlos según las funciones discriminantes resultantes del resto de casos (Tabachnick & Fidell, 2001). En este trabajo hemos utilizado ambos modelos, como se verá a continuación.

2. Resultados

Para lograr el primer objetivo (comparar la capacidad discriminante de los distintos parámetros acústicos relacionados con las pausas llenas: formantes y parámetros relacionados con la cualidad de voz), se han evaluado los resultados del ALD incorporando al análisis, en primer lugar, únicamente las variables relacionadas con la estructura de formantes de las pausas llenas; en segundo lugar, solo las variables referentes a la cualidad de voz; y finalmente, todas las variables conjuntamente.

Los diagramas de dispersión de los Gráficos 1 a 3 muestran cómo la simple combinación de los valores de los tres primeros formantes evidencia un grado destacable de solapamiento entre los distintos hablantes. Si bien en algunos casos todos los valores aparecen en un espacio más bien compacto (por poner un ejemplo, en los hablantes 3 y 4), en otros los valores están muy dispersos en el gráfico (por ejemplo, los hablantes 1, 7 o 10). Estos gráficos se complementan con los datos de la Tabla 4, que muestran los valores medios y la desviación estándar para los tres formantes de las distintas pausas llenas.

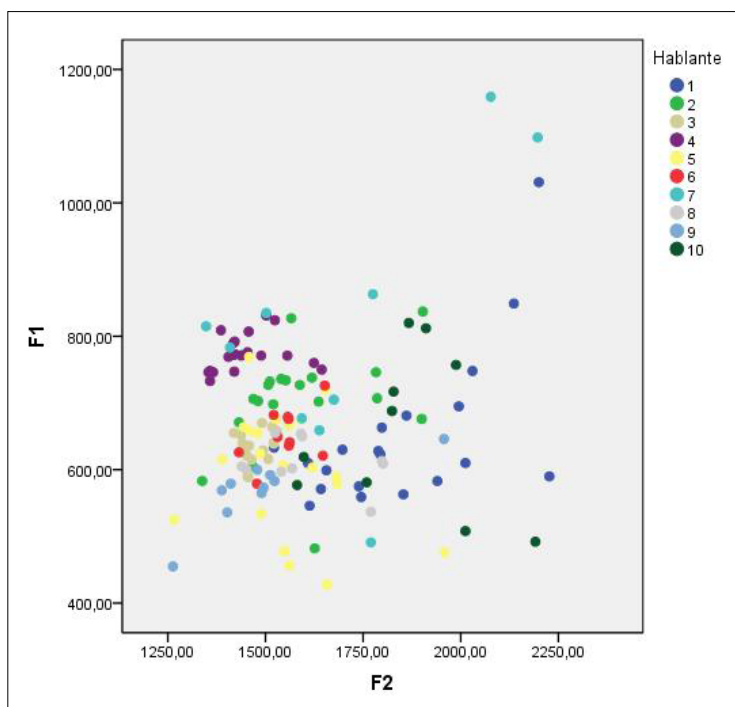


Gráfico 1. Diagrama de dispersión con los valores de F1 y F2.

Tabla 4. Valores medios de los formantes y su desviación estándar. Aparecen sombreados los valores de desviación estándar superiores a la media, que se concentran mayoritariamente en los hablantes 1, 7 y 10.

Hablante	media F1	media F2	media F3	Desv. estándar F1	Desv. estándar F2	Desv. estándar F3
1	649,35	1842,9	2800,65	25,74	45,66	39,07
2	705,45	1583,25	2856,65	17,75	34,33	12,12
3	635,85	1479,9	2592,2	5,84	8,78	23,98
4	774,3	1450,7	2974,75	6,1	18,7	26,79
5	599,85	1552,95	2638,25	20,59	31,54	54,22
6	651,5	1549,9	2567,9	12,92	21,21	53,41
7	808,5	1698,4	2608,9	63,44	85,85	25,18
8	612,2	1607,8	2813,9	11,16	42,96	16,04
9	569,8	1492	2682,2	15,57	57,32	36,41
10	657,1	1855,9	2883,1	37,74	58,74	39,22
media	666,39	1611,37	2741,85	21,685	40,509	32,644

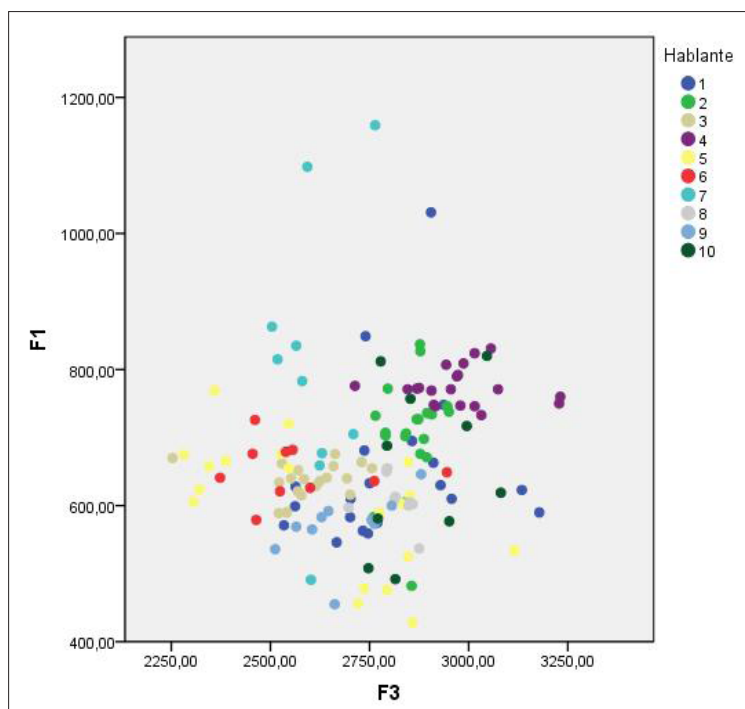


Gráfico 2. Diagrama de dispersión con los valores de F1 y F3.

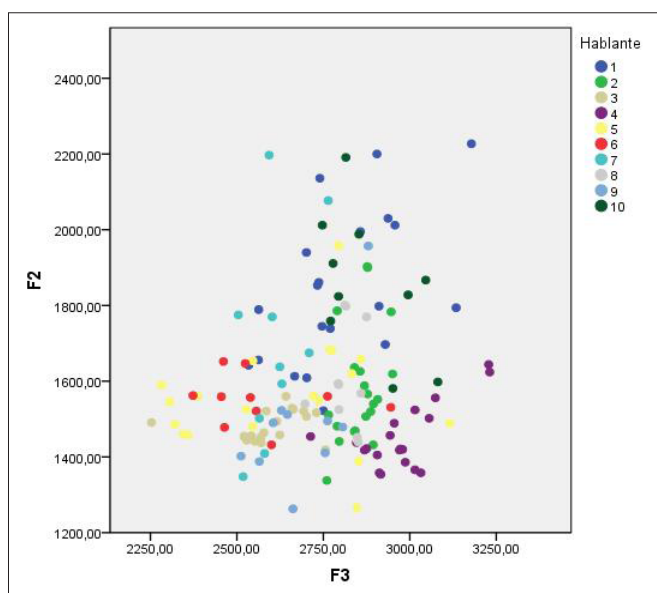


Gráfico 3. Diagrama de dispersión con los valores de F2 y F3.

Mediante el análisis de las dos primeras funciones discriminantes (Gráfico 4) se continúa observando un alto grado de solapamiento entre los hablantes (solapamiento que se vería solventado en parte con la representación gráfica de las demás funciones discriminantes). En los gráficos siguientes, los centroides indican los puntos medios de las funciones discriminantes para cada uno de los hablantes.

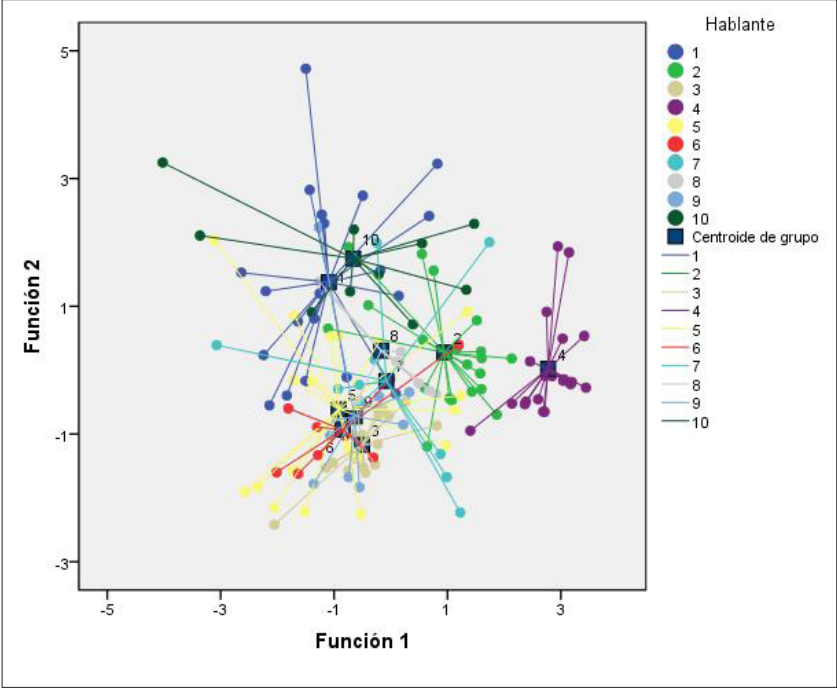


Gráfico 4. Diagrama de dispersión con los valores correspondientes a las dos primeras funciones discriminantes y los centroides para cada hablante para el ALD mediante las variables referidas a formantes.

En la misma línea, el Gráfico 5 muestra las dos primeras funciones discriminantes del ALD teniendo en cuenta las variables referentes a la cualidad de voz. Se observa una mejor zonificación de los hablantes, aunque persisten ciertos solapamientos que se verían reducidos si se graficaran las otras funciones discriminantes.

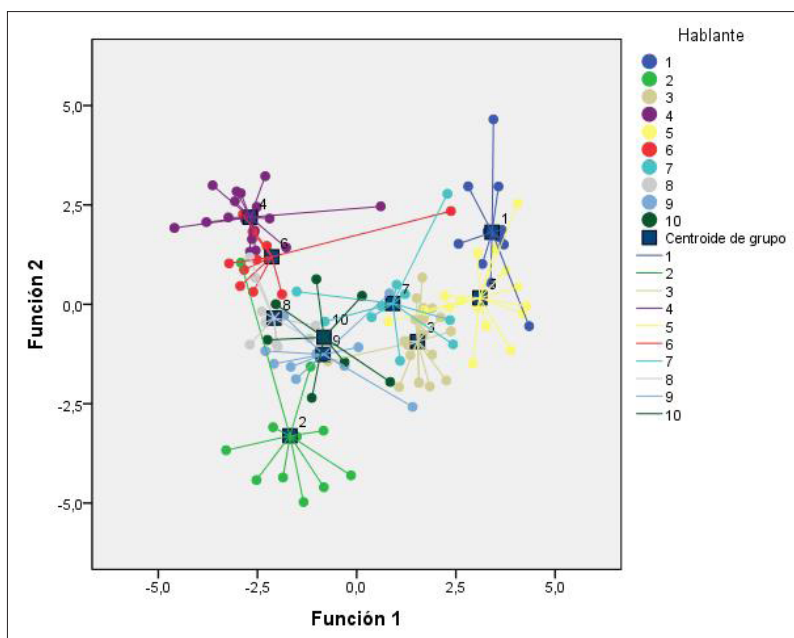


Gráfico 5. Diagrama de dispersión con los valores correspondientes a las dos primeras funciones discriminantes y los centroides para cada hablante para el ALD mediante las variables referidas a la cualidad de voz.

Finalmente, mediante el Gráfico 6 se observa como el tratamiento integrado de las variables referentes a los formantes y a la cualidad de voz mejora considerablemente los resultados. En los Gráficos 4 a 6 se observa como progresivamente disminuye el solapamiento de los casos, lo que implica un mayor porcentaje de casos clasificados correctamente (Tabla 5), tanto en la clasificación de los casos originales (es decir, del conjunto de casos utilizados para calcular las funciones discriminantes) como en la validación cruzada.

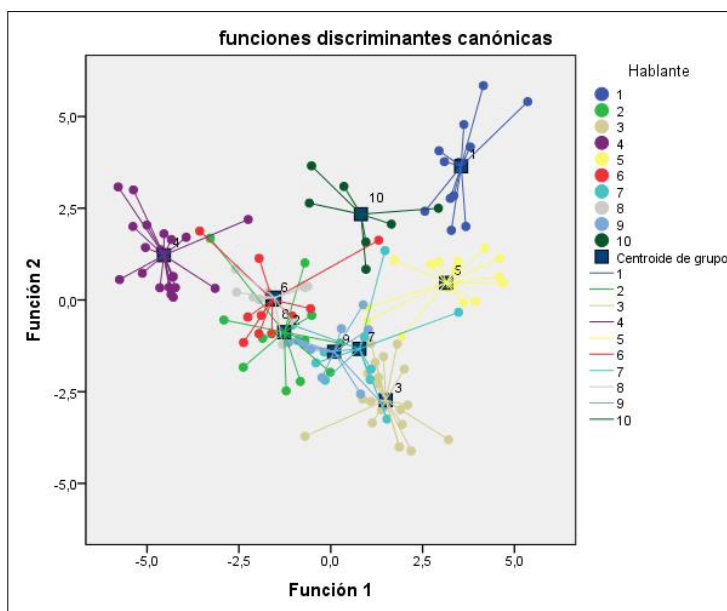


Gráfico 6. Diagrama de dispersión con los valores correspondientes a las dos primeras funciones discriminantes y los centroides para cada hablante para el ALD mediante todas las variables (referidas a formantes y a la cualidad de voz).

Tabla 5. Resumen del porcentaje de casos clasificados correctamente según las variables analizadas.

	Casos originales clasificados correctamente	Casos clasificados correctamente en la validación cruzada
Análisis de formantes: F1, F2, F3, $\Delta F2-F1$, $\Delta F3-F1$, $\Delta F3-F2$	51%	47,7%
Análisis de la cualidad de voz: mediana de F0, valor mínimo de F0, porcentaje de segmentos sonoros, <i>jitter</i> absoluto, <i>shimmer</i> absoluto, razón ruido-armónicos	79,5%	62,2%
Análisis conjunto: formantes y cualidad de voz	87,2%	68,5%

En síntesis, los resultados obtenidos utilizando únicamente las variables relacionadas con los formantes lleva a un porcentaje de clasificaciones correctas discreto (47,7% en la validación cruzada), similar al obtenido por Foulkes et al. (2004). Este porcentaje, aunque bajo, multiplica por 5 la probabilidad debida al azar. Los resultados mejoran a medida que se introducen más variables (las relacionadas con la cualidad de voz) en el análisis (hasta el 68,5%), lo que permite validar la primera de las hipótesis de trabajo.

Por otro lado, considerando la totalidad de variables, pero introduciendo en el modelo del ALD únicamente a los hablantes masculinos, se obtiene un porcentaje de

casos originales agrupados correctamente ligeramente inferior (83,8%) en comparación con el ALD de todos los hablantes juntos, pero el acierto en la validación cruzada es 5 puntos mejor (73,8%). Tomando en consideración solo a los hablantes femeninos, los resultados son mejores en ambas clasificaciones: 91,3% de clasificaciones correctas de los casos originales, y 81,2% en la validación cruzada. La mejoría en los resultados viene determinada por dos factores: el primero, y más obvio, es que al reducir a la mitad el número de hablantes introducidos en el modelo, la probabilidad de acierto por azar se duplica; el segundo viene determinado por el papel de la variable F0. Cuando se analizan los hablantes masculinos y femeninos conjuntamente, el F0 es una variable que discrimina eficazmente entre hombres y mujeres, pero es menos útil para discriminar entre hablantes concretos. Al analizar únicamente hablantes del mismo sexo, el F0 revela diferencias más sutiles entre hablantes.

Sin embargo, aunque el análisis discriminante es una herramienta útil para evaluar el poder discriminante de las variables lingüísticas objeto de estudio (Grant, 2007; Spassova, 2009; López, 2010; Turell, 2010), en la práctica forense se deben tener en cuenta básicamente dos aspectos metodológicos. En primer lugar, el uso de esta técnica es especialmente indicada cuando se trabaja con un corpus indubitado de tipo cerrado, y no con un corpus de tipo abierto. Y en segundo lugar, como el perito trabaja habitualmente con una única muestra dubitada (o grupo de muestras dubitadas) que debe comparar con un grupo de muestras indubitadas, tiene que decidir a cuál de los hablantes que conforman el corpus indubitado se corresponde la muestra dubitada analizada. Para comprobar la capacidad discriminante bajo estas premisas (segundo objetivo), se ha repetido el proceso tomando como corpus indubitado las grabaciones indubitadas, y se han utilizado las grabaciones dubitadas como casos nuevos. En estas circunstancias, el ALD ha atribuido correctamente una media del 90% de pausas llenas dubitadas al hablante correcto (Tabla 6), lo que permite validar también la segunda hipótesis. En la tabla se especifica, para las identificaciones incorrectas, a qué hablante se ha atribuido erróneamente la pausa llena.

Tabla 6. Resumen del porcentaje de casos clasificados correctamente según las variables analizadas.

Hablante	Identificaciones correctas	Identificaciones incorrectas. Entre paréntesis, el hablante al que erróneamente se ha atribuido la pausa llena
1 (hombre)	80%	20% (5)
2 (mujer)	80%	20% (8, 10)
3 (hombre)	100%	
4 (mujer)	100%	
5 (hombre)	90%	10% (7)
Media	90%	10%

CONCLUSIONES

A raíz de los resultados obtenidos, se puede concluir que el análisis de las pausas llenas produce resultados interesantes para la comparación forense de grabaciones. En este estudio hemos evaluado su capacidad discriminante comparando los porcentajes de acierto en las clasificaciones utilizando la información relativa a su estructura de formantes, a un conjunto de variables relacionadas con la cualidad de voz, y al conjunto de variables (relacionadas tanto con los formantes como con la cualidad de voz). Los resultados (expresados en porcentaje de casos clasificados correctamente) han resultado mejores en el análisis de la cualidad de voz que en el análisis de los formantes. Asimismo, con la consideración conjunta de las variables se han obtenido aún mejores resultados. Así, respecto a la primera hipótesis, según la cual la inclusión en el análisis discriminante de las variables relacionadas con la cualidad de voz mejoraría los resultados respecto a los obtenidos mediante el análisis de formantes, hay que señalar que los resultados obtenidos han permitido validarla. Los porcentajes de clasificaciones correctas en la validación cruzada han pasado del 47,7% cuando se utilizaban únicamente las variables referentes a los formantes, al 62,2% cuando las variables utilizadas fueron las de la cualidad de voz, y finalmente al 68,5% cuando se utilizaron todas las variables conjuntamente. Estos porcentajes son aceptables, teniendo en cuenta que las probabilidades de clasificación correctas debidas al azar son del 10%. Respecto a la segunda hipótesis, que predecía que se obtendrían porcentajes altos de pausas llenas clasificadas correctamente mediante el análisis discriminante y un corpus indubitado de tipo cerrado, los datos de este estudio han alcanzado una media del 90% de acierto.

En síntesis, este estudio confirma que las pausas rellanas, en la línea de los estudios de Künzel (1997), Foulkes et al. (2004), Cicres y Turell (2005) y Cicres (2007), tienen capacidad discriminante y, consecuentemente, utilidad en la comparación forense de voces. La principal aportación del presente estudio es que demuestra que con el análisis conjunto de variables relativas a la estructura de formantes y a la cualidad de voz se obtienen mejores resultados en la identificación de hablantes mediante el ALD que únicamente con el análisis de uno u otro grupo de variables. Con el análisis conjunto se obtienen porcentajes medios de clasificaciones correctas del 90% en las pruebas con un solo grupo de pausas llenas dubitadas y un corpus indubitado de 10 hablantes.

Sin embargo, asumimos en este estudio ciertas limitaciones metodológicas. En primer lugar, el hecho de que el ALD es apropiado si se utiliza un corpus indubitado de tipo cerrado. En segundo lugar, las grabaciones dubitadas e indubitadas analizadas tienen que tener una calidad de sonido aceptable y con similares características acústicas. Finalmente, hay que tener en cuenta que las variables referentes a la cualidad de voz son muy sensibles a factores externos tales como el estado de salud,

las emociones, el nivel de alcohol, etc. (Johnstone & Scherer, 1999; Hollien & Martin, 1996). En cualquier caso, la comparación forense de voces es un trabajo complejo que requiere del análisis de múltiples aspectos de la voz y del uso lingüístico de los hablantes (Rose, 2002; Kredens & Goráewska-Lach, 1998; Langeveld-Cambier, 2007; Nolan, 2001).

REFERENCIAS BIBLIOGRÁFICAS

- Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, 5(9/10), 341-345.
- Brown, R. (1982). What is speaker recognition. *Journal of the International Phonetics Association*, 12(1), 13-24.
- Burridge, D. & Mulder, J. (1988). *English in Australia and New Zealand: An introduction to its structure, history and use*. Melbourne: Oxford University Press.
- Cicres, J. (2007). Análisis discriminante de un conjunto de parámetros fonéticoacústicos de las pausas llenas para identificar hablantes. *Síntesis Tecnológica*, 3(2), 87-98.
- Cicres, J. & Turell, M. T. (2005). El análisis multidimensional de la voz como herramienta para la identificación del hablante en fonética forense. En P. Cano (Coord.), *Actas del VI Congreso de Lingüística General* (pp. 803-812). Santiago de Compostela: Servicio de Publicaciones de la Universidad de Santiago de Compostela.
- Coulthard, M. (1994). On the use of corpora in the analysis of forensic texts. *Forensic Linguistics. The International Journal of Speech Language and the Law*, 1(1), 27-43.
- Coulthard, M. (2004). Author identification, idiolect, and linguistic uniqueness. *Applied Linguistics*, 25(4), 431-447.
- Delgado, C. (2005). Comentarios sobre el contexto actual de la identificación forense de locutores. En M. T. Turell (Ed.), *Lingüística forense, lengua y derecho: conceptos, métodos y aplicaciones* (pp. 113-129). Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra, Documenta Universitaria.
- Erickson, F. (1979). Talking down: Some cultural sources of miscommunication in interracial interviews. En A. Wolfgang (Ed.), *Nonverbal Behavior: Applications and Cultural Implications* (pp. 99-126). Nueva York: Academic Press.
- Foulkes, P., Carrol, G. & Hughes, S. (2004). *Sociolinguistic and acoustic variability in filled pauses*. Ponencia presentada en la International Association for Forensic Phonetics and Acoustics Conference, Helsinki, Universidad de Helsinki.
- French, J. P. & Harrison, P. (2007). Position statement concerning use of impressionistic likelihood terms in forensic speaker comparison cases. *International Journal of Speech, Language and the Law*, 14(1), 137-144.
- French, J. P., Nolan, F., Foulkes, P., Harrison, P. & McDougall, K. (2010). The UK position statement on forensic speaker comparison: A rejoinder to Rose and Morrison. *The International Journal of Speech Language and the Law*, 17(1), 143-152.

- Gibbons, J. & Turell, M. T. (Eds.) (2008). *Dimensions of Forensic Linguistics*. Ámsterdam/ Filadelfia: John Benjamins.
- Gold, E. & French, P. (2011). International practices in forensic speaker comparison. *International Journal of Speech, Language and the Law*, 18(2), 293-307.
- González, J., Cervera, T. & Miralles, J. L. (2002). Análisis acústico de la voz: Fiabilidad de un conjunto de parámetros multidimensionales. *Acta Otorrinolaringológica Española*, 53, 256-268.
- González, J. & Lucena, J. J. (2005). IDENTIVOX©: Un sistema automático de reconocimiento de locutores por la voz para acústica forense Lengua y Derecho. En M. T. Turell (Ed.), *Lingüística forense, lengua y derecho: Conceptos, métodos y aplicaciones* (pp. 131-143). Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra, Documenta Universitaria.
- Grant, T. (2007). Quantifying evidence for forensic authorship analysis. *International Journal of Speech, Language and the Law*, 14(1), 1-25.
- Halliday, M. A. K. (1989). *Language, context and text. Aspects of language in a social-semiotic perspective*. Oxford: Oxford University Press.
- Hollien, H. (2002). *Forensic voice identification*. Londres: Academic Press Forensics.
- Hollien, H. & Martin, C.A. (1996). Conducting research on the effects of intoxication on speech. Forensic Linguistics. *The International Journal of Speech, Language and the Law*, 3(1), 107-128.
- Johnstone, T. & Scherer, K. R. (1999). *The effects of emotions on voice quality*. Ponencia presentada en el 14th International Conference of Phonetic Sciences, San Francisco.
- Kienast, M. & Glitza, F. (2003). Respiratory sounds as an idiosyncratic feature in speaker recognition. En *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 1607-1610). Bellaterra: Universitat Autònoma de Barcelona.
- Kredens, K. & Gorálewska-Lach, G. (1998). Language as sole incriminating evidence: The Augustynek case. Forensic Linguistics. *The International Journal of Speech, Language and the Law*, 5(2), 193-202.
- Künzel, H. J. (1997). Some general phonetic and forensic aspects of speaking tempo. Forensic Linguistics. *The International Journal of Speech, Language and the Law*, 4(1), 48-83.
- Langeveld-Cambier, T. (2007). Current methods in forensic speaker identification: Results of a collaborative exercise. *The International Journal of Speech, Language and the Law*, 14(2), 223-243.

- López, F. (2010). *El análisis de las características dinámicas de la señal de habla como posible marca para la comparación e identificación forense de voz: Un estudio para el español de México*. Tesis doctoral, Universitat Pompeu Fabra, Barcelona, España.
- Mulac, A., Erlandson, K.T., Farrar, W. J., Hallett, J. S., Molloy, J. L. & Prescott, M. E. (1998). 'Uh-huh. What's that all about?': Differing Interpretations of Conversational Backchannels and Questions as Sources of Miscommunication Across Gender Boundaries. *Communication Research*, 25(6), 641-668.
- Nolan, F. (1983). *The phonetic bases of speaker recognition*. Cambridge: Cambridge University Press.
- Nolan, F. (2001). Speaker identification evidence: Its forms, limitations, and roles. *En Proceedings of the Conference Law and Language: Prospect and retrospect*, Levi (Laponia finesa).
- Pätzold, M. & Simpson, A. (1995). An acoustic analysis of hesitation particles in German. *En Proceedings of the 13th International Congress of Phonetic Sciences* (pp. 512-515). Stockholm.
- Rose, P. (2002). *Forensic speaker identification*. Londres: Taylor & Francis.
- Rose, P. & Morrison, G. (2009). A response to the UK position statement on forensic speaker comparison. *The International Journal of Speech, Language and the Law*, 16, 139-193.
- Spasova, M. S. (2009). *El potencial discriminatorio de las secuencias de categorías gramaticales en la atribución forense de autoría de textos en español*. Tesis doctoral, Universitat Pompeu Fabra, Barcelona.
- Tabachnick, B. G. & Fidell, L. S. (2001). *Using multivariate statistics*. Londres: Allyn & Bacon.
- Tannen, D. (1990). *You just don't understand: Men and women in conversation*. Nueva York: William Morrow.
- Turell, M. T. (2010). The use of textual, grammatical and sociolinguistic evidence in forensic text comparison. *The International Journal of Speech, Language and the Law*, 17(2), 211-250.
- Ward, N. (2006). Non-Lexical conversational sounds in American English. *Pragmatics and Cognition*, 14(1), 113-184.