

Hacia la aplicabilidad de la cualidad de la voz en fonética judicial¹

Marianela Fernández Trinidad

Universidad Complutense de Madrid
marianela.fernandez@ucm.es ORCID: <https://orcid.org/0000-0002-0087-0829>

Enviado: 03/09/2022; Aceptado: 30/01/2023; Publicado en línea: 05/07/2023

Cómo citar este artículo / Citation: Marianela Fernández Trinidad (2022). Hacia la aplicabilidad de la cualidad de la voz en fonética judicial. *Loquens*, 9(1-2), e093, <https://doi.org/10.3989/loquens.2022.e093>.

RESUMEN: La cualidad de la voz derivada de ajustes laríngeos a largo plazo se erige como una característica potencialmente individualizadora de los hablantes. Esto la coloca en una situación ventajosa respecto de otros parámetros fonéticos utilizados en lingüística forense. Sin embargo, cualquiera que se enfrente a su análisis tropezará enseguida con una dificultad metodológica derivada de su naturaleza intrínsecamente multidimensional. En ello radica su principal desventaja y la razón fundamental por la que su análisis no siempre se considere en la aproximación tradicional utilizada en la comparación de hablantes con fines de identificación en el contexto de un peritaje. Este trabajo muestra, a partir de un estudio experimental sobre el disimulo de la voz mediante falsete, que es posible trabajar con un conjunto reducido de parámetros laríngeos responsables de la cualidad de la voz y, a la vez, facilitar su interpretación y explicación, algo fundamental en el ámbito pericial de la lingüística forense.

Palabras clave: Disimulo; cualidad laríngea de la voz; fonética forense; aplicabilidad.

ABSTRACT: *Towards the applicability of voice quality in forensic phonetics.* - Voice quality derived from long-term laryngeal settings stands out as a potentially individualizing trait of speakers. This places it in an advantageous situation with respect to other phonetic parameters used in forensic linguistics. However, anyone confronted with its analysis will immediately run into a methodological difficulty stemming from its inherently multidimensional nature. In this lies its main disadvantage and the fundamental reason why its analysis is not always considered in the traditional approach used in the comparison of speakers for identification purposes. Based on an experimental inquiry on voice disguised by means of falsetto, this study shows that it is possible to work with a reduced set of laryngeal features responsible for voice quality and facilitate its interpretation and explanation, which is a critical issue for forensic practice.

Key Words: Voice disguise; laryngeal voice quality; forensic phonetics; applicability.

¹ El presente trabajo retoma algunos de los principales resultados obtenidos de diferentes análisis acústicos realizados en Fernández Trinidad (2018, inédito) para ser interpretados y discutidos a propósito de su relevancia y aplicabilidad en fonética judicial. En este sentido, agradezco muy especialmente a los dos revisores anónimos por sus comentarios y observaciones y, también, al profesor Gómez-Vilda por sus sugerencias.

1. INTRODUCCIÓN Y ANTECEDENTES

El presente estudio se enmarca en el ámbito de las ciencias del habla con aplicaciones forenses y persigue dos finalidades. De una parte, profundiza en el estudio de la voz disimulada con falsete para detallar los parámetros acústicos y sus correlatos articulatorios que lo diferencian del modo habitual de fonación conocido como *modal voice* (en adelante, modal o habitual, indistintamente). De otra parte, busca probar que es posible reducir a un número limitado de parámetros, independientes e interpretables, el análisis de la cualidad de la voz para que su estudio pudiera aplicarse en un peritaje.

La elección de los parámetros que se considerarán en una comparación forense de hablantes está sujeta a las características de las muestras dubitadas e indubitadas disponibles en términos de cantidad, calidad y comparabilidad (cf., entre otros, Gil Fernández *et al.*, 2017, y Rose, 2002). De acuerdo con Nolan (1983), Rose (2002) y Wolf (1972), los parámetros elegidos deberían cumplir los siguientes requisitos:

- Alta frecuencia de aparición en la muestra.
- Fácil extracción, medición e interpretación.
- Independientes.
- Poca variación dentro del mismo sujeto (variabilidad intralocutor baja) y mayor variación entre sujetos distintos (variabilidad interlocutor alta).
- Resistentes a la transformación o al disimulo.

El parámetro fonético ineludible, a pesar de que no suele ser en sí mismo concluyente, es la f_0 . Su valor medio, junto con los valores de dispersión y distribución a ella asociados (mínimo, máximo, desviación típica, etc.) se analizan y comparan de forma sistemática en un cotejo de voces. Es evidente que, ante una sospecha de disimulo mediante el aumento drástico de la f_0 estos valores resultarán difícilmente comparables y, por tanto, de escasa utilidad para el cotejo. Es así como se vuelve todavía más necesario profundizar en el estudio de otros parámetros laríngeos responsables de la cualidad de voz que cumplan, además, con el mayor número de requisitos recomendados en fonética judicial.

El estudio de la voz en falsete es relevante en el contexto judicial porque, al suponer una distorsión considerable con respecto a la voz habitual de una persona, resulta eficaz para enmascararla con propósitos delictivos. El disimulo puede afectar tanto a las muestras dubitadas de grabaciones que se obtienen de la comisión de un delito, como a las indubitadas recogidas de los sospechosos en sedes judiciales o policiales. Se ha comprobado su uso en casos de secuestros, extorsiones y amenazas (cf. por ejemplo, Zhang y Tan, 2008). Asimismo, en el momento de recoger muestras de voces indubitadas los sospechosos podrían intentar alterar o disimular su forma de hablar aumentando su tono habitual de voz². En el momento de cometer el delito, es-

pecialmente si este es premeditado, podrían preferirse los procedimientos tecnológicos de transformación de la voz. Actualmente, es posible acceder de forma sencilla a variadas tecnologías y aplicaciones que cambian el tono de la voz, le añaden efectos y distorsiones, imitan otras voces, etc. Sin embargo, durante las declaraciones que servirán para el análisis de las muestras de voces indubitadas, el acusado solamente podría recurrir a algún procedimiento humano (no electrónico) de disimulo del habla, siguiendo la división propuesta por Rodman (1998).

Varios estudios (Figueiredo y de Souza Britto, 1996; Künzel, 2000; Masthoff, 1996; Perrot y Chollet, 2008; Praveena y Krishna, 2015) apuntan a que las alteraciones de la cualidad de voz derivadas de ajustes laríngeos exclusivamente, como es el caso de la voz de falsete o del susurro (*whisper*), son formas bastante frecuentes de disimulo. Las razones principales radican en que con este tipo de alteraciones resulta fácil producir y mantener una voz impostada; asimismo, no interfieren con la transmisión del contenido del mensaje, y son eficaces porque logran enmascarar la voz original del hablante, consiguiendo despistar a los expertos del ámbito forense y a los sistemas automáticos de reconocimiento de locutores (Perrot, Aversano y Chollet, 2007; Perrot, *et al.* 2007; Perrot y Chollet, 2008; Künzel, González-Rodríguez y Ortega-García, 2004; Zhang y Tan, 2008).

Además de que el falsete ha interesado mayormente como registro de la voz cantada y no hablada, muchos de los parámetros fonéticos que hasta ahora se han medido con técnicas invasivas no son factibles en el contexto forense, donde solo podemos apoyarnos en el análisis acústico y perceptivo. En ese sentido, siguen faltando estudios que corroboren de una manera fehaciente las correlaciones entre las características acústicas y las bases fisiológicas de la voz de falsete como mecanismo de disimulo del habla.

2. OBJETIVOS Y MÉTODO

2.1. Objetivos

El trabajo persigue dos objetivos, como se ha adelantado en la introducción:

1. Profundizar en los rasgos acústicos que caracterizan la fonación disimulada en falsete (en contraste con la voz habitual) y en sus correlatos o bases fisiológicas.
2. Poner a prueba el estudio multiparamétrico de la fonación mediante un procedimiento que, en caso de que se cumpla el requisito de cantidad de las muestras (Rose, 2002), pudiera aplicarse en un peritaje del ámbito forense para analizar la cualidad de la voz derivada de ajustes laríngeos a largo plazo.

En cuanto al estudio acústico de los rasgos laríngeos de la voz, interesa aislar, por un lado, los parámetros caracte-

² En la nota periodística “La voz, prueba contra el crimen machista” publicada por *El País* (06/11/15), expertos de la policía científica explicaban: “A la hora de tomar una muestra de voz de un acusado, los guardias civiles y policías se enfrentan a todo tipo de estratagemas,

desde las más obvias [negarse a hacerlo, lo que también puede resultar indicativo para el juez] a las más retorcidas simulaciones y enmascaramientos del acento o la forma de hablar.” https://elpais.com/tecnologia/2015/11/04/actualidad/1446663289_487919.html

rizadores del habla en falsete y, por otro, aquellos que podrían llegar a ser individualizadores de las voces. La idea que sustenta este planteamiento es la existencia de rasgos laríngeos idiosincrásicos de los hablantes que son muy difíciles o imposibles de alterar mediante procedimientos humanos de disimulo vocal (cf. Alves *et al.*, 2012, 2014; Fernández Trinidad y Rojo, 2019). De una parte, esperamos poder distinguir un cierto número de rasgos que, por ser intrínsecos al cambio de voz hacia el falsete, haya tenido que modificar cualquier locutor que disimule su voz y, de otra parte, otros rasgos que no se hayan visto alterados. Los primeros permitirían caracterizar el habla en falsete, mientras que los segundos podrían contribuir a la caracterización de las distintas voces y, por extensión, también de los individuos.

El segundo objetivo de este trabajo es analizar el comportamiento laríngeo mediante un procedimiento factible de ponerse en práctica en fonética judicial. En este sentido, tal procedimiento debería cumplir los siguientes requisitos:

- a. Que las voces se puedan medir y analizar mediante una técnica sencilla y no invasiva. En el contexto forense, solo podrá ser a partir de la señal acústica.
- b. Que los parámetros acústicos que permiten describir las voces sean independientes, interpretables y explicables en términos acústicos y fisiológicos.

2.2. Protocolo de grabación

El corpus utilizado fue recogido en el contexto de un estudio más amplio (cf. Fernández Trinidad, 2018). Participaron 6 locutores masculinos italianos de Nápoles, con estudios universitarios concluidos y edades comprendidas entre los 29 y 34 años ($M = 31.67$). Antes de las grabaciones, los informantes completaron sus datos personales y declararon no tener patologías vocales. Fueron debidamente instruidos acerca de los objetivos del estudio, así como del procedimiento experimental, y firmaron un consentimiento informado. Las sesiones de grabación se realizaron en la cámara insonorizada del Laboratorio de Fonética del Consejo Superior de Investigaciones Científicas (CSIC) de Madrid. Se utilizó un micrófono de condensador omnidireccional Earset AE E6i (Contryman), una interfaz de audio AEUA-25EX (Ronald) y el software Adobe Audition (Microsoft). La frecuencia de muestreo fue de 44100 Hz, con una resolución de 16 bits. A los participantes se les pidió que grabaran una misma frase de 11 sílabas, primero con su voz habitual y, luego, con una voz simulada en falsete. La razón principal que justifica la elección fue conseguir frases idénticas en su composición segmental para centrar el estudio acústico (y otro perceptivo reportado en Fernández Trinidad y Rojo, 2019) en la cualidad de la voz derivada de ajustes laríngeos a largo plazo. De otra parte, en la práctica forense se recomienda obtener del imputado muestras indubitadas de su habla espontánea (en caso de que este tenga una actitud colaborativa) y también de la lectura de frases o textos (cf. por ejemplo, Gil, Alves y Hierro, 2012 y Gil Fernández *et al.*, 2017).

Se buscó que el enunciado estuviera formado por sílabas con la estructura CV, donde C fuera sonora y no nasal, y V fuera una [a] tónica. Esta decisión metodológica es habitual entre los estudios que analizan los parámetros laríngeos responsables de la cualidad de la voz. Por ejemplo, Palacios (2018)³ explica que conviene descartar vocales altas y cerradas o excesivamente nasalizadas y escoger la [a] o la [e] porque aseguran una mejor reconstrucción de la fuente glótica a partir de la señal acústica.

Se consiguió para cada locutor un total de 60 repeticiones válidas de la misma frase, 30 para cada registro fonatorio.

2.3. Extracción de parámetros laríngeos

De las frases grabadas se segmentaron con Praat 30 vocales [a] tónicas por registro de fonación (15 en voz modal y 15 en voz de falsete) para cada uno de los 6 locutores, consiguiéndose así un total de 180 muestras para su posterior análisis. Se calculó el valor medio de los parámetros laríngeos de forma automática utilizando el programa Biomet®Phon [versión 2.3] de BiometroSoft®. Este programa reconstruye la onda glotal a partir de la señal acústica mediante un filtrado inverso que resta la función de transferencia ejercida por el tracto vocal y el efecto de la radiación labial. A partir de ello, estima el valor de 72 parámetros relativos al comportamiento laríngeo. Los parámetros referidos a la fuente glótica extraídos se reproducen en la Tabla 1 a partir de BioMetroSoft® (2014). Puede consultarse una descripción de los rasgos glotales en Gómez-Vilda *et al.* (2013), Palacios (2018), Palacios *et al.* (2020), Rodellar-Biarge *et al.* (2015). Según se explica en estos mismos estudios, los parámetros laríngeos que están relacionados entre sí forman distintos grupos: a) f_0 y perturbaciones asociadas a ella, b) cepstrales, c) perfil espectral, d) parámetros biomecánicos de los pliegues vocales, e) parámetros de la onda glótica de base temporal, f) parámetros de defecto de cierre glótico, y g) temblor.

3. RESULTADOS DEL ANÁLISIS CONVENCIONAL

3.1. Comportamiento de la f_0

Comenzamos presentando los resultados del comportamiento de la f_0 en el cambio de voz hacia el falsete por ser, en este tipo de modificación voluntaria de la voz, el cambio más drástico desde un punto de vista perceptivo y acústico, y por tratarse del parámetro más utilizado en las comparaciones forenses de hablantes, como se ha explicado en §1.

³ “Esta interacción [entre la fuente y el filtro] se basa en la reverberación de la onda por su paso por los pliegues vocales hasta llegar a la zona de radiación [los labios]. Las vocales cerradas como son la /i/ y la /u/, además de algunas consonantes nasales, ofrecen una mayor alteración en el patrón de fonación de la fuente glótica.” (Palacios, 2018, p. 64). También Gobl y Ni Chasaide (2010, p. 380).

Tabla 1. Lista de los 72 parámetros relativos al comportamiento glótico que estima *Biomet®Phon*.

Parámetros	Parámetros
Absolute Pitch	<i>Body Stiffness</i>
<i>Absolute Normal Jitter</i>	<i>Body Mass Unbalance</i>
<i>Absolute Normal Area Shimmer</i>	<i>Body Losses Unbalance</i>
<i>Absolute Normal Minimum Sharpness</i>	<i>BodyStiffnessUnbalance</i>
<i>Noise-HarmonicRatio, NHR</i>	<i>Cover Mass</i>
<i>Muc. /AvAc. Energy, MAE</i>	<i>Cover Losses</i>
<i>MWC Cepstral 1</i>	<i>Cover Stiffness</i>
<i>MWC Cepstral 2</i>	<i>Cover Mass Unbalance</i>
<i>MWC Cepstral 3</i>	<i>Cover Losses Unbalance</i>
<i>MWC Cepstral 4</i>	<i>Cover Stiffness Unbalance</i>
<i>MWC Cepstral 5</i>	<i>Rel. Recovery 1 Time</i>
<i>MWC Cepstral 6</i>	<i>Rel. Recovery 2 Time</i>
<i>MWC Cepstral 7</i>	<i>Rel. Open 1 Time</i>
<i>MWC Cepstral 8</i>	<i>Rel. Open 2 Time</i>
<i>MWC Cepstral 9</i>	<i>Rel. Max. Amplitude Time</i>
<i>MWC Cepstral 10</i>	<i>Rel. Recov. 1 Ampl.</i>
<i>MWC Cepstral 11</i>	<i>Rel. Recov. 2 Ampl.</i>
<i>MWC Cepstral 12</i>	<i>Rel. Open 1 Ampl.</i>
<i>MWC Cepstral 13</i>	<i>Rel. Open 2 Ampl.</i>
<i>MWC Cepstral 14</i>	<i>Rel. Stop Flow Time</i>
<i>MW PSD 1st Max. ABS.</i>	<i>Rel. Start Flow Time</i>
<i>MW PSD 1st Min. rel.</i>	<i>Rel. Closing Time</i>
<i>MW PSD 2nd Max. rel.</i>	<i>Val. Flow GAP</i>
<i>MW PSD 2nd Min. rel.</i>	<i>Val. Contact GAP</i>
<i>MW PSD 3rd Max. rel.</i>	<i>Val. Adduction GAP</i>
<i>MW PSD End Val. rel.</i>	<i>Val. Permanent GAP</i>
<i>MW PSD 1st Max. Pos. ABS.</i>	<i>1st. Order Cyclic Coefficient</i>
<i>MW PSD 1st Min. Pos. rel.</i>	<i>2nd. Order Cyclic Coefficient</i>
<i>MW PSD 2nd Max. Pos. rel.</i>	<i>3rd. Order Cyclic Coefficient</i>
<i>MW PSD 2nd Min. Pos. rel.</i>	<i>Physiological Tremor Freq.</i>
<i>MW PSD 3rd Max. Pos. rel.</i>	<i>Physiological Tremor Ampl.</i>
<i>MW PSD End Val. Pos. rel.</i>	<i>Neurological Tremor Freq.</i>
<i>MW PSD 1st Min NSF</i>	<i>Neurological Tremor Ampl.</i>
<i>MW PSD 2nd Min NSF</i>	<i>Fluttering Tremor Freq.</i>
<i>Body Mass</i>	<i>Fluttering Tremor Ampl.</i>
<i>Body Losses</i>	<i>Tremor amplitude (rMSA)</i>

Tabla 2: Máximo, mínimo, media y desviación estándar (SD) de f_0 (Hz) en voz modal para cada locutor y el promedio entre sujetos.

	Modal						
	L1	L2	L3	L4	L5	L6	Promedio
Máx.	150.1	119.1	120.4	117.2	162.2	153.5	137.1
Media	129.7	110.1	103.5	110.4	121.8	124.1	116.6
SD	12.9	8.3	9.6	6.0	23.4	15.9	12.7
Mín.	113.8	96.3	90.2	100.6	94.4	108.6	100.7

Tabla 3: Máximo, mínimo, media y desviación estándar (SD) de f_0 (Hz) en falsete para cada locutor y el promedio entre sujetos.

	Falsete						
	L1	L2	L3	L4	L5	L6	Promedio
Máx.	355.9	374.3	326.9	433.7	441.4	519.7	408.6
Media	306.4	302.8	284.1	378.3	366.3	336.4	329.0
SD	49.4	63.7	42.2	30.1	52.4	123.2	60.2
Min.	145.7	172.4	153.0	319.1	281.7	218.6	216.1

Tabla 4: Diferencias de f_0 entre falsete y voz modal (por locutor y promedio entre sujetos). Las diferencias se expresan en hercios, octavas y cents.

	Falsete - Modal						
	L1	L2	L3	L4	L5	L6	Promedio
Hercios	176.6	192.6	180.6	267.8	244.5	212.2	212.4
Octavas	1.24	1.46	1.46	1.78	1.59	1.44	1.50
Cents	1488	1751	1748	2132	1906	1726	1792

Las Tablas 2 y 3 ofrecen los valores máximos, mínimos y medios de la f_0 alcanzados en voz modal y falsete por cada locutor y también el promedio general obtenido en los 6 locutores. La Tabla 4 resume las diferencias alcanzadas en los valores de f_0 entre falsete y voz habitual para cada locutor y el promedio global obtenido de los 6 locutores analizados.

Los valores máximos promedio fueron de 137.1 Hz (modal) y 408.6 Hz (falsete), mientras que los mínimos registrados fueron 100 Hz (modal) y 216 Hz (falsete). La f_0 media observada se ubicó en torno a los 116 Hz (modal) y 329 Hz (falsete), mientras que la desviación típica promedio fue de 12.7 y 60.2 Hz para la voz modal y en falsete, respectivamente. Por tanto, se pudo comprobar que, cuando modifican su voz hacia el falsete, los locutores analizados aumentan su f_0 de promedio 1,5 octavas (212.4 Hz) respecto de su tono de voz habitual y que la desviación típica media es mayor en la voz fingida que en la habitual. La mayor variabilidad en el habla fingida se ha comprobado con otros tipos de disimulo, como era de prever (cf. por ejemplo, Alves *et al.* 2012; Lahoz-Bengoechea, Villa y Gil Fernández, 2017, entre muchos otros que podrían mencionarse).

3.2. Comportamiento de los 72 rasgos glotales

La Tabla 5 recoge los rasgos laríngeos que fueron modificados por cada locutor al cambiar de voz habitual a falsete y la Tabla 6 la proporción media de parámetros glotales alterados en el cambio de voz, según los resultados de un test *t* (corrección Bonferroni). En ambas tablas los rasgos se organizan por grupos de parámetros (§2.3) y en la Tabla 6 aparecen ordenados de mayor a menor variación.

Con una proporción media superior a 0.9, los parámetros cepstrales (grupo B) son los que resultan más afectados

Tabla 5: Número de rasgos laríngeos modificados ($p < 0.05$) por locutor en el cambio de voz habitual a falsete organizados por grupo de parámetros.

	total	L1	L2	L3	L4	L5	L6
A	6	3	4	3	4	6	3
B	14	13	12	10	14	14	14
C	14	9	8	11	12	11	6
D	12	12	7	7	12	10	6
E	12	9	4	8	12	11	3
F	4	1	2	2	3	2	2
G	10	5	1	5	4	4	1
total	72	52	38	46	61	58	35

Tabla 6: Proporción media de rasgos laríngeos modificados ($p < 0.05$) como consecuencia del cambio de voz (modal-falsete), organizados por grupos de parámetros y ordenados de mayor a menor variación. SD: desviación estándar; SE: error típico.

Parámetros	Proporción	SD	SE
B	0.92	0.11	0.05
D	0.75	0.22	0.09
C	0.68	0.16	0.07
E	0.65	0.30	0.12
A	0.64	0.19	0.08
F	0.50	0.16	0.06
G	0.33	0.19	0.08

cuando un sujeto cambia de su voz habitual al falsete. A estos les siguen los rasgos responsables de reflejar la biomecánica de las cuerdas vocales (grupo D) con una proporción media de 0.75. A continuación, encontramos los parámetros que informan sobre el perfil espectral de las voces (grupo C), los que dan cuenta de la base temporal de la onda glótica (grupo E) y los parámetros relacionados con la perturbación de la f_0 (grupo A), con valores medios de 0.68, 0.65 y 0.64, respectivamente. Los grupos que presentan la menor variación recogen información sobre el defecto de cierre glótico (grupo F), de los que varía la mitad de los parámetros (0.5), y, finalmente, los parámetros responsables del temblor (grupo G), de los cuales varía solo un tercio (0.33).

Los resultados hasta aquí comentados ya permiten una caracterización a nivel glotal el cambio de voz hacia el falsete. Sin embargo, considerando que en este tipo de análisis se maneja un número muy elevado de parámetros y que muchos de ellos están correlacionados entre sí, decidimos realizar un Análisis por Componentes Principales (o CPA por sus siglas en inglés).

4. PROPUESTA DE ANÁLISIS POR COMPONENTES PRINCIPALES

4.1. Comportamiento de los CPs

El elevado número de rasgos glotales sumado al alto nivel de redundancia informativa como consecuencia de la interdependencia de los parámetros dentro de cada grupo dificultan tanto el análisis como la interpretación de los datos. Recuérdese, además, que en las comparaciones forenses de voces se desaconseja expresamente utilizar parámetros interdependientes (cf. §1). Para solventar las dificultades señaladas de cara a abordar el análisis de la cualidad de la voz en un contexto forense proponemos llevar a cabo un análisis factorial por Componentes Principales (en adelante, CPs). Este tipo de análisis es apropiado porque consigue, al mismo tiempo, reducir la dimensionalidad de conjuntos complejos de datos cuya distribución aún no se conoce, y hacerlos más fácilmente interpretables (cf. Jolliffe, 1986, Godino, Gómez-Vilda y Blanco, 2006).

Apoyándonos en los resultados obtenidos del análisis efectuado por grupos de parámetros (§3.2), simplificamos el número original de rasgos para el análisis factorial por CPs. Al prescindir de los parámetros cepstrales (grupo B), los de perfil espectral (grupo C) y los relativos al temblor (grupo G), el número de variables originales se reduce de 72 a 32. La exclusión se fundamenta en los resultados del análisis previo. Se prescinde de los parámetros de temblor (grupo G) por no ser pertinentes para el análisis que nos ocupa. A fin de facilitar la interpretación también se excluyen del análisis factorial los parámetros que no tienen un valor claro hasta el momento (cepstrales, grupo B) o que no tienen un significado sólido o fiable cuando se da un cambio de fonación hacia la voz de falsete, como ocurre con los parámetros de perfil espectral (grupo C)⁴.

⁴ La información ofrecida por el espectro armónico en el falsete es difícil de tratar porque la reconstrucción de la envolvente espectral requiere de un alto nivel de interpolación o inferencia entre los datos muestrales.

El análisis CPA genera, a partir de operaciones matemáticas, nuevas variables o CPs que, al comienzo, coinciden exactamente con el número de variables originales, en nuestro caso, los 32 parámetros laríngeos. Luego, y puesto que no todos los parámetros tienen la misma relevancia explicativa, se conservan, por convención, los que son capaces de explicar al menos un 5 % de la varianza total de los datos (cf. Baayen, 2008). Una vez aplicado el criterio umbral del 5 % para seleccionar los parámetros o componentes con mayor poder explicativo, las 32 variables originales quedaron agrupadas en 4 CPs que juntos explican el 81,75 % de la distribución total de los datos. Recuérdese que entre los CPs no existen nunca solapamientos explicativos puesto que son independientes. Por tanto, y como se observa en la Tabla 7, el CP1 explica el 27,30 % de la varianza de los datos, el CP2 explica un 21,63 % adicional, el CP3 añade un 17,83 % y el CP4 añade un 14,98 % hasta alcanzar el 81,75 % de varianza explicada. En la Tabla 8 se ofrece la distribución de los 32 rasgos laríngeos en los 4 CPs y el grado de correlación (positiva o negativa) entre cada rasgo original y el CP que lo agrupa.

Como ilustra la Tabla 8, los 32 rasgos laríngeos se han asociado con 4 CPs y, dentro de cada CP, los parámetros quedaron ordenados por valor absoluto de correlación según un coeficiente de signo positivo o negativo. Así, podemos ver que los rasgos originales *Absolute pitch* y *Body mass* se correlacionan fuertemente con el primer componente principal (CP1), aunque el *pitch* presenta una correlación positiva (.936) y la masa del cuerpo de los pliegues vocales mantiene una correlación negativa con el CP1 (-.772). Esto significa que, ambos rasgos están fuertemente correlacionados con el CP1 que los agrupa y que, cuando el valor del CP1 aumenta, se incrementa el *pitch* y disminuye la masa de los pliegues vocales implicada en la vibración.

Para comprender qué rasgos laríngeos son realmente relevantes en la configuración de cada uno de los cuatro CPs, se aislaron aquellos parámetros con valor absoluto de correlación superior a 0.7, pues observamos a partir de los datos que no hay ningún parámetro con un valor superior en varios componentes principales (CPs).

Las tablas 9-12 reúnen los rasgos o parámetros más relevantes que componen los cuatro CPs, ordenados de mayor a menor según su valor absoluto de correlación.

El CP1 agrupa 9 de los rasgos glotales originales y recoge información tonal, pues se centra en la f_0 y en los ajustes laríngeos que permiten aumentar o disminuir el tono. A excepción de la masa del cuerpo de los pliegues vocales (*Body mass*), el resto de los rasgos se relaciona positivamente con el *pitch*, es decir, aumentan su valor cuando aumenta la f_0 .

El CP2 aglutina 6 rasgos relacionados con instantes temporales de la onda glotal, pues todos ellos dan cuenta de los momentos temporales relevantes durante un ciclo de fonación. La correlación de ellos con el CP2 es positiva.

Si el CP2 se asociaba con la dimensión temporal de la onda glotal, el CP3 recoge rasgos que dan cuenta de la dimensión de las amplitudes correspondientes. Los cuatro rasgos más relevantes en la definición de este CP mantienen una relación negativa con él.

Tabla 7: Extracción de los 4 CPs ordenados según el porcentaje (%) de explicación de varianza (de mayor a menor).

CPs	% varianza	% acumulado
1	27.307	27.307
2	21.631	48.937
3	17.838	66.775
4	14.982	81.757

Tabla 8: Distribución de los 32 rasgos glotales en los 4 CPs con sus grados de correlación positiva o negativa. Las casillas en blanco no alcanzaron una correlación con valor absoluto mínimo de .300.

	CPs			
	1	2	3	4
Absolute pitch	.936			
Body stiffness	.895			
Noise Harmonic Ratio	.894			
Cover losses	.874		-.305	
Cover stiffness	.856			
Cover mass unbalance	.818			
Cover losses unbalance	.793			.442
Cover stiffness unbalance	.786			
Body mass	-.772			
Rel. stop flow time	.622	.468	.463	
Val. flow GAP	-.620	.547		
Rel. open 2 time		.967		
Rel. max. ampl. time		.960		
Rel. closing time		.932		
Rel. open 1 time		.923		
Rel. start flow time	-.437	.834		
Rel. recov. 2 time	.392	.824		
Val. permanent GAP	.493	-.638		
Rel. recov. 1 time	.577	.628		
Rel. open 1 ampl.			-.916	
Rel. recov. 2 ampl.			-.883	
Rel. recov. 1 ampl.			-.822	
Rel. open 2 ampl.		-.501	-.809	
Val. adduction GAP			.796	
Cover mass			-.668	-.354
Abs. norm. shimmer	-.489		.540	.358
Val. contact GAP		.303	.399	
Body stiffness unbalance				.946
Body mass unbalance				.921
Abs. norm. jitter			.379	.861
Body losses		.359		-.797
Body losses unbalance			.470	.748

Tabla 9: Rasgos laríngenos originales agrupados en el CP1 por valor absoluto de correlación.

CP1	
Parámetros	Coefficiente factorial
Absolute pitch	.936
Body stiffness	.895
Noise Harmonic Ratio	.894
Cover losses	.874
Cover stiffness	.856
Cover mass unbalance	.818
Cover losses unbalance	.793
Cover stiffness unbalance	.786
Body mass	-.772

Tabla 10: Rasgos laríngenos originales agrupados en el CP2 por valor absoluto de correlación.

CP2	
Parámetros	Coefficiente factorial
Rel. open 2 time	.967
Rel. max. ampl. time	.960
Rel. closing time	.932
Rel. open 1 time	.923
Rel. start flow time	.834
Rel. recov. 2 time	.824

Tabla 11: Rasgos laríngenos originales agrupados en el CP3 por valor absoluto de correlación.

CP3	
Parámetros	Coefficiente factorial
Rel. open 1 amplitud	-.916
Rel. recov. 2 amplitud	-.883
Rel. recov. 1 amplitud	-.822
Rel. open 2 amplitud	-.809
Val. adduction GAP	.796

Tabla 12: Rasgos laríngenos originales agrupados en el CP4 por valor absoluto de correlación.

CP4	
Parámetros	Coefficiente factorial
Body stiffness unbalance	.946
Body mass unbalance	.921
Abs. norm. jitter	.861
Body losses	-.797
Body losses unbalance	.748

Finalmente, el CP4 reúne rasgos laríngenos vinculados con el comportamiento del cuerpo de los pliegues vocales e incluye también el *jitter*, que da cuenta de la variabilidad temporal de la f_0 . A excepción de las pérdidas de energía (*Body losses*), todos los rasgos presentan una correlación positiva.

Un test de medias con prueba de Bonferroni demostró que el CP1 es el único diferenciador en el cambio de voz (habitual-falsete) para todos los locutores, como muestran los datos recogidos en las Tablas 13 y 14.

Tabla 13: Contraste de medias de los 4 CPs en el cambio de voz (falsete-modal) para todos los locutores. Los valores de la misma fila con distinto subíndice (a, b) son significativamente distintos en $p < 0.05$.

	Falsete	Modal
CP1	.90363 _a	-.90363 _b
CP2	.05062 _a	-.05062 _a
CP3	.08756 _a	-.08756 _a
CP4	-.05015 _a	.05015 _a

Tabla 14: Contraste de medias de los 4 CP en el cambio de voz (habitual-falsete) por locutor (L). Los valores de la misma fila con distinto subíndice (a, b) son significativamente distintos en $p < 0.05$.

		Falsete	Modal
L1	CP1	.54137 _a	-.69249 _b
	CP2	-.17152 _a	-2.02209 _b
	CP3	.30977 _a	-.06369 _a
	CP4	.99664 _a	-.22832 _b
L2	CP1	.83115 _a	-.80410 _b
	CP2	.03742 _a	-.22940 _a
	CP3	-.10373 _a	1.55220 _b
	CP4	.64472 _a	.28434 _a
L3	CP1	.77149 _a	-1.32498 _b
	CP2	.31729 _a	.30246 _a
	CP3	.25842 _a	-.69912 _b
	CP4	-1.02627 _a	-.56898 _b
L4	CP1	1.47695 _a	-1.19954 _b
	CP2	.18501 _a	1.46173 _b
	CP3	-.36066 _a	-1.03654 _b
	CP4	-.02930 _a	-.73767 _a
L5	CP1	1.41513 _a	-.66571 _b
	CP2	.22089 _a	-.96776 _b
	CP3	-.81648 _a	-1.08473 _b
	CP4	-.45736 _a	.47007 _b
L6	CP1	.38567 _a	-.73495 _b
	CP2	-.28536 _a	1.15135 _b
	CP3	1.23807 _a	.80650 _a
	CP4	-.42930 _a	1.08144 _b

4.2. Rasgos de la voz disimulada en falsete

El primer objetivo de este estudio (§2.1) consistía en caracterizar en términos acústicos y fisiológicos el habla disimulada en falsete en contraste con la voz habitual. Además, se buscaba hacerlo exclusivamente a partir de las muestras de voz grabadas.

El CP1 resultó ser el componente discriminatorio en el cambio de voz habitual a falsete, pues todos los rasgos laríngeos que lo definen variaron significativamente para cada uno de los locutores sin excepción. Como se ha visto, el CP1 se centra en el *pitch*, correlato perceptivo de la f_0 , y en el comportamiento biomecánico de las cuerdas vocales, en especial, en el de la cubierta, de acuerdo con descripciones anteriores (por ejemplo, Titze, 1994 y Núñez, 2013).

La tensión soportada por la cubierta de las cuerdas vocales (*Cover stiffness*) mantiene una correlación alta y positiva (.856) con el *pitch*, cuyos valores aumentan significativamente en el paso hacia la voz de falsete. Esto provoca grandes gastos de energía e irregularidad en la vibración de la cubierta de los pliegues, como demuestra la proyección positiva de los parámetros *Cover losses* (.874), *Cover mass unbalance* (.818), *Cover stiffness unbalance* (.786) y *Cover losses unbalance* (.793).

De otra parte, los rasgos incluidos en el CP1 que dan cuenta del comportamiento no ya de la cubierta, sino del cuerpo de los pliegues vocales, son *Body stiffness* y *Body mass*. Estos dos parámetros parecen influir directamente en el aumento (o en su caso, disminución) de la f_0 y lo hacen de forma inversa. La tensión del cuerpo de las cuerdas vocales contribuye positivamente en el aumento del *pitch* con una proyección de .895, al tiempo que la masa del cuerpo de las cuerdas vocales implicadas en la vibración lo hace de forma negativa o inversamente proporcional al aumento del *pitch*, con una proyección negativa de -.772.

Finalmente, del análisis efectuado por componentes principales, se observa que el parámetro que da cuenta de la proporción entre ruido y armonicidad (*Noise Harmonic Ratio*) presenta valores más altos en la voz disimulada en falsete que en la fonación habitual de todos los locutores analizados. Este parámetro, índice de turbulencia, podría ser un buen indicador de defectos de cierre en la glotis.

4.3. Calidad individual de la voz

Como se desprende de los datos relativos a los otros tres CPs presentados en las Tablas 13 y 14, los parámetros glotales agrupados bajo los CP2, CP3 y CP4 muestran una mayor variación entre sujetos. A diferencia de lo que ocurre con el CP1, estos componentes presentan un comportamiento fluctuante, unas veces cambiaron y otras no, en función del locutor.

El primer objetivo de este estudio se dirigía a poder aislar el grupo de rasgos laríngeos que, por ser intrínsecos al cambio de fonación (habitual-falsete), todos los sujetos tuvieran que alterar para conseguir el disimulo buscado y que han resultado ser los parámetros reunidos en el CP1.

Como consecuencia de haber podido aislar estos rasgos fonéticos definitorios del falsete (CP1) han quedado también mejor acotados el grupo de parámetros vocales invariantes, aquellos que en la muestra analizada no se han visto modificados siempre por todos los locutores analizados en el cambio a falsete. Algunos de estos rasgos podrían ser potencialmente discriminantes, aunque difícilmente pudiera tratarse del mismo rasgo para todas las personas.

4.4. Utilidad práctica del análisis en fonética judicial

En la práctica forense partimos de que es habitual encontrarse con que la información de que se dispone es escasa y ha sido recogida por canales telefónicos o microfónicos limitantes, en ocasiones en medio de ambientes ruidosos, lo que da como resultado muestras sonoras de baja calidad. Esta realidad forense cuestiona la aplicabilidad del análisis acústico y estadístico al margen de la herramienta que vaya a utilizarse.

El problema de la cantidad es inapelable, pues siempre se necesitará un tamaño mínimo de datos suficiente para establecer comparaciones estadísticas fiables. Sin embargo, la dificultad que supone la baja calidad de las muestras para la fonética forense podría mitigarse con la utilización de una herramienta que analice la función glótica (cf. Gómez-Vilda, 2022). El estudio multiparamétrico de la cualidad de la voz mediante el análisis del comportamiento glótico ofrece ciertas ventajas sustanciales que es necesario señalar, especialmente en casos de disponer de material suficiente, pero no disponer de material sonoro de óptima calidad. En primer lugar, la mala calidad de la señal resulta más sensible al análisis de la función articulatoria que al de la función glótica, puesto que son precisamente las zonas de altas frecuencias las que se verán más afectadas. La información glótica, en la medida en que se basa en el análisis de cada ciclo de la fuente glótica en el dominio del tiempo, siendo esta una señal obtenida mediante la atenuación de los efectos amplificadores de las resonancias del tracto vocal, resulta menos sensible al deterioro de altas frecuencias, por lo que resiste mejor las limitaciones habituales de la práctica forense relacionadas con la calidad de las muestras disponibles para el análisis (conversación personal con el Dr. Gómez-Vilda). En segundo término, el análisis de la cualidad de voz resultante de configuraciones laríngeas a largo plazo, además de tener un alto poder discriminante, como ya se ha comentado, presenta otro valor añadido en el que es preciso insistir: el comportamiento glótico resiste mejor los intentos de impostación, camuflaje, o disimulo. Como señalan los estudios realizados sobre distintos mecanismos humanos de disimulo, el funcionamiento laríngeo resulta más difícil de modificar o imitar, pues no parece que tengamos un control tan preciso y consciente de nuestro sistema fonador, en comparación al que ejercemos sobre el articulador (cf. por ejemplo, Alves *et al.*, 2012, 2014; Fernández Trinidad y Rojo, 2019, y los trabajos recogidos en Lahoz-Bengoechea, Villa y Gil Fernández, 2017).

En la comparación forense de voces sigue siendo de utilidad el método denominado “mixto” o “híbrido”, es de-

cir, la complementariedad del análisis acústico-perceptivo “clásico” o “tradicional” llevado a cabo por expertos fonetistas con el análisis automático (por ejemplo, Drygajlo *et al.*, 2015; Gold y French, 2011). Es innegable la eficiencia de los sistemas automáticos que, hoy en día, alcanzan una prestación con altos grados de precisión, especialmente si las muestras son suficientemente comparables. Sin embargo, como contrapartida negativa, son más difícilmente interpretables, pues cada vez está más alejada la posibilidad de comprender -y, en consecuencia, explicar- el modo en que los algoritmos funcionan realmente⁵ (Gunning *et al.*, 2019). Además, la capacidad analítica de los especialistas continúa siendo, hasta el momento, irremplazable.

Este trabajo demuestra, por un lado, que un análisis que dé cuenta de la función glótica, como el que ofrece la herramienta BioMet®Soft, resulta provechoso para el estudio de la fonación, puesto que ha permitido observar con precisión el comportamiento laríngeo de los locutores a partir de las muestras de audio. Los valores correspondientes a los parámetros glotales fueron fácil y cómodamente extraíbles a partir de la señal acústica.

De otra parte, del análisis con CPA también se derivan resultados prometedores, pues se ha demostrado que es posible abordar el estudio de la cualidad de voz, en concreto de los parámetros que dan cuenta del comportamiento laríngeo, disminuyendo el número de factores a unos pocos CPs, independientes, interpretables y, por tanto, explicables.

Estos son aspectos decisivos, pues varios estudios (por ejemplo, González-Rodríguez *et al.* 2014; Jessen, 1997; San Segundo y Gómez-Vilda, 2017, entre otros) han señalado la importancia de considerar la cualidad de voz laríngea en el ámbito de la fonética forense por su potencial discriminante. Sin embargo, como se ha explicado, por ejemplo, en Gil Fernández y San Segundo, (2014), su estudio se excluye a menudo por suponer una máxima complejidad de análisis y de interpretación.

El análisis tradicional realizado por fonetistas, apoyado en una herramienta como la que se ha utilizado para este estudio, y complementado con un análisis por CPA, demuestra que es viable reducir los rasgos laríngeos que caracterizan la fonación a unos pocos factores independientes entre sí e interpretables.

5. CONCLUSIONES

Luego de este estudio se conocen mejor cuáles son los factores o parámetros que parecen variar de forma sistemática en todos los locutores cuando cambian su voz hacia el falsete. Estos serían los parámetros englobados en el CP1 que se centran en la f_0 y en los ajustes laríngeos responsables de su aumento; principalmente, el comportamiento de la cubierta de los pliegues vocales.

⁵ En este punto conviene recordar (§3.2) el comportamiento de los parámetros cepstrales de la onda glotal, los rasgos que mayor variación experimentaron en el cambio de voz analizado. El estudio pormenorizado de este grupo podría ser pertinente para comprender el peor rendimiento que los reconocedores automáticos tienen con las voces disimuladas, incluso cuando han sido previamente entrenados con una base de datos apropiada (Künzel, González-Rodríguez y Ortega García, 2004).

Los análisis llevados a cabo confirman, por tanto, que hay una serie de rasgos glotales que cambian sistemáticamente y para todos los locutores cuando modifican la voz habitual hacia el falsete. Estos parámetros, englobados en el CP1, dependen del tipo de fonación y no del locutor.

En la bibliografía (recuérdese, Titze, 1994 o Núñez 2013) ya se había señalado que en el caso del falsete se daba un aumento drástico de la f_0 y que la responsabilidad de ese aumento recaía fundamentalmente en el comportamiento de la cubierta del pliegue vocal. Lo que no estaba claro, sin embargo, era si, acompañando a estos cambios, otros parámetros también se modificarían de forma sistemática y significativa. La respuesta es negativa, con base en las pruebas obtenidas en este estudio. Se ha visto que solo algunos locutores cambian – y de forma asistemática – ciertos parámetros laríngeos, agrupados en los CP2, CP3 y CP4, pero que no todos lo hacen. Por tanto, entre esos invariantes podrían hallarse algunos potencialmente asociados a la individualidad del locutor. Los rasgos clave habrá que determinarlos en cada caso, pues los mismos parámetros no funcionarán, presumiblemente, para todos los hablantes. Por último, conviene recordar que estos resultados están, como es lógico, condicionados a la muestra analizada y, por tanto, resulta imprescindible seguir explorando el comportamiento de los rasgos laríngeos y su posible relevancia forense con un conjunto de datos más extenso y variado.

Otro resultado interesante de este estudio indica que es posible encarar el análisis de la cualidad laríngea de la voz (que se ha visto tiene un alto poder discriminante) a pesar de la conocida desventaja que supone su multidimensionalidad. El análisis propuesto demuestra que es posible simplificar el elevado número de rasgos glotales responsables de la cualidad de la voz a unos pocos CPs independientes e interpretables en términos acústicos y también fisiológicos. Ganancias explicativas como estas contribuyen a facilitar la comunicación de un análisis fonético complejo al ámbito judicial.

Siempre en el terreno aplicado de la fonética forense o judicial, conviene señalar la utilidad de los programas informáticos semiautomáticos para el análisis de la fonación. La utilización de este tipo de herramientas no invasivas, por supuesto, tiene el valor adicional de permitir conocer el comportamiento glótico a partir de una señal exclusivamente acústica, único método viable en un cotejo de voces con fines judiciales.

No obstante, el análisis propuesto en este trabajo necesitará ulteriores comprobaciones que terminen de validarlo. En este sentido, conviene recordar que, de acuerdo con las directrices de referencia como son la *European Network of Forensic Science Institutes* (ENFSI, 2015), la *International Association for Forensic and Legal Linguistics* (IAFL) o la *International Association for Forensic Phonetics and Acoustics* (IAFPA, 2020), cualquier método de análisis debería contar con la validación y aceptación previa de la comunidad científica para poder ser utilizado en evaluaciones periciales.

Vale la pena seguir dedicando esfuerzos en el camino hacia la aplicabilidad del análisis de la cualidad de la voz

conducentes a encontrar metodologías y herramientas que permitan resolver las principales dificultades que plantea el contexto forense.

6. REFERENCIAS

- Alves, H., Fernández Trinidad, M., Gil Fernández, J., Infante, P., Lahoz-Bengoechea, J. M., Pérez Sanz, C. y San Segundo, E. (2012). Disguised voices: A perceptual experiment. *3rd European Conference of the International Association of Forensic Linguistics (IAFL)*, Porto.
- Alves, H., Gil Fernández, J., Pérez Sanz, C. y San Segundo, E. (2014). La cualidad individual de la voz y la identificación del locutor: el proyecto CIVIL. En Y. Congosto, M. L. Montero Curiel, y A. Salvador Plans (Eds.), *Fonética experimental, educación superior e investigación* (Vol. 1, pp. 591-612). Madrid: Arco/Libros.
- Baayen, R. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics*. Cambridge: Cambridge University Press.
- BioMetroSoft®. (2022). *Informe técnico para documentar la funcionalidad de la aplicación informática BioMet®Fore* [versión 1.0].
- BioMetroSoft®. (2014). *BioMet®Phon: Tool for de Evaluation of Voice Quality and Biometri*. User's Manual [versión 2.3].
- Drygajlo, A., Jessen, M., Gfroerer, S., Wagner, I., Vermeulen, J. y Niemi, T. (2015). *Methodological Guidelines for Best Practice in Forensic Semiautomatic and Automatic Speaker Recognition*. Frankfurt: Verlag für Polizeiwissenschaft. https://enfsi.eu/wp-content/uploads/2016/09/guidelines_fasr_and_fasr_0.pdf
- ENFSI, European Network of Forensic Science Institutes (2015). *ENFSI Guideline or Evaluative Reporting in Forensic Science*, http://enfsi.eu/wp-content/uploads/2016/09/ml_guideline.pdf
- Fernández Trinidad, M. (2018). *Caracterización del falseto y sus consecuencias para la discriminación de voces*. [Tesis doctoral inédita]. Universidad Nacional de Educación a Distancia.
- Fernández Trinidad, M. y Rojo, J. (2019). Perceptual cues for individual voice quality. En J. Gil Fernández y M. Gibson (Eds.), *Romance Phonetics and Phonology* (pp. 161-176). Oxford University Press.
- Figueiredo, R. y de Souza Britto, H. (1996). A report on the acoustic effects of one type of disguise. *Forensic Linguistics*, 3, 168-175. <http://dx.doi.org/10.1558/ijsl.v3i1.168>
- Gil, J. Alves, H. y Hierro, J. A. (2012). Proposition raisonnée de protocole de capture de voix connue à des fins judiciaires. *Revue Internationale de Criminalistique et de Police Scientifique et Technique*, LXV, 319-345.
- Gil Fernández, J., Fernández Trinidad, M., Infante, P. y Lahoz-Bengoechea, J. M. (2017). "Obtaining speech samples for research and expertise in forensic phonetics". En: Orletti, F. y Mariottini, L. (Eds.) *Theories, Practices, Instruments of Forensic Linguistics* (pp. 27-50). Cambridge Scholars Publishing.
- Gil Fernández, J. y San Segundo, E. (2014). La cualidad de voz en fonética judicial. En E. Garayzábal, M. Jiménez y M. Reigosa (Coords.), *Lingüística Forense. La Lingüística en el ámbito legal y policial* (pp. 154 -199). Madrid: Euphonía Ediciones.
- Gobl, C. y Ni Chasaide, A. (2010). Voice source variation and its communicative functions. En W. Hardcastle, J. Laver, y F. Gibbon (Eds.), *The Handbook of Phonetic Sciences* (2.ª ed., pp. 378-423). Oxford: Wiley-Blackwell.
- Godino, J., Gómez-Vilda, P. y Blanco, M. (2006). Dimensionality reduction of a pathological voice quality assessment system based on Gaussian mixture models and short-term cepstral parameters. *IEEE Transactions on Biomedical Engineering*, 53(10), 1943-1953.
- Gold, E. y French, P. (2011). International practices in forensic speaker comparison. *International Journal of Speech, Language, and the Law*, 18(2), 293-307. <http://dx.doi.org/10.1558/ijsl.v18i2.293>
- Gómez-Vilda, P. (2022). *Informe técnico para documentar la funcionalidad de la aplicación informática BioMet®Fore*. Versión 1.0, 2022 NeuSpeLab, Glottex VAS (www.glottalso-lutions.com)

- Gómez-Vilda, P.; Rodellar Biarge, M. V.; Nieto Lluís, V.; Martínez Olalla, R.; Álvarez Marquina, A.; Scola Yurrita, B.; Ramírez Calvo, C.; Poletti Serafini, D. y Fernández Fernández, M. (2013). BioMet@Phon: A system to monitor phonation quality in the clinics. En *eTELEMED 2013: The Fifth International Conference on eHealth, Telemedicine, and Social Medicine* (pp. 253-258), Nice, France. ISBN 978-1-61208-252-3.
- González-Rodríguez, J., Gil, J., Pérez, R. y Franco-Pedroso, J. (2014). What are we missing with i-vectors? A perceptual analysis of i-vector based falsely accepted trials. *Proceedings of Odyssey 14. The Speaker and Language Recognition Workshop* (pp. 33-40).
- Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S. y Yang, G-Z. (2019). XAI- Explainable artificial intelligence. *Science Robotics*, 4(37), eaay7120. <http://dx.doi.org/10.1126/scirobotics.aay7120>
- IAFPA, International Association of Forensic Phonetics and Acoustics (2020). *Code of Practice*. <http://www.iafpa.net/wp-content/uploads/2020/09/IAFPA-Code-of-Practice-2020.pdf>
- IAFL, International Association of Forensic and Legal Linguistics (s/d). *Code of Practice*. https://www.iafl.org/wp-content/uploads/2018/07/IAFL_Code_of_Practice_1-1.pdf
- Jessen, M. (1997). Speaker-specific information in voice quality parameters. *The International Journal of Speech, Language, and the Law*, 4(1), 84-103. <http://dx.doi.org/10.1558/ijsl.v4i1.84>
- Jolliffe, I. (1986). Principal component analysis and factor analysis. En *Principal Component Analysis* (pp. 115-128). New York: Springer.
- Künzel, H. (2000). Effects of voice disguise on speaking fundamental frequency. *Forensic Linguistics*, 7, 149-179.
- Künzel, H., González-Rodríguez, J. y Ortega-García, J. (2004). Effect of voice disguise on the performance of a forensic automatic speaker recognition system. En *Proceedings of Odyssey 04. The Speaker and Language Recognition Workshop*. (pp. 1-4).
- Lahoz-Bengoechea, J. M., Villa Villa, J. y Gil Fernández, J. (2017). Fillers in disguised accented speech. *13th Biennial Conference of the International Association of Forensic Linguists*.
- Masthoff, H. (1996). A report on a voice disguise experiment. *Forensic Linguistics*, 3(1), 160-167. <http://dx.doi.org/10.1558/ijsl.v3i1.160>
- Nolan, F. (1983). *The Phonetic Bases of Speaker Recognition*. Cambridge University Press.
- Núñez, F. (2013). Fisiología de la fonación. En I. Cobeta, F. Núñez y S. Fernández (Eds.), *Patología de la voz* (pp. 55-75). Barcelona: Marge Medica Books.
- Palacios, D. (2018). *Contribución al estudio de selección de parámetros para identificación de estrés en la voz*. [Tesis doctoral inédita]. Universidad Politécnica de Madrid.
- Palacios, D., Rodellar, V., Lázaro, C., Gómez, A. y Gómez, P. (2020). An ICA-based method for stress classification from voice samples. *Neural Computing and Applications*, 32(24), 17887-17897. <http://dx.doi.org/10.1007/s00521-019-04549-3>
- Perrot, P., Aversano, G. y Chollet, G. (2007). Voice disguise and automatic detection: review and perspectives. En Y. Stylianou, M. Faundez-Zanuy y A. Esposito (Eds.), *Progress in Nonlinear Speech Processing* (pp. 101-117). Berlin: Springer.
- Perrot, P. y Chollet, G. (2008). The question of disguised voice. *The Journal of the Acoustical Society of America*, 123(5), 3878. <http://dx.doi.org/10.1121/1.2935782>
- Perrot, P., Preteux, C., Vasseur, S. y Chollet, G. (2007). Detection and recognition of voice disguise. *Proceedings IAFPA 2007* (pp. 1-3). Plymouth, UK: The College of St Mark y St John.
- Praveena, J. y Krishna, Y. (2015). Identifying speaker from disguised speech using aural perception and Mel-frequency cepstral coefficient. *Journal of Indian Speech Language & Hearing Association*, 29(2), 28-34. <http://dx.doi.org/10.4103/0974-2131.185974>
- Rodellar-Biarge, V., Palacios-Alonso, D., Nieto-Lluís, V. y Gómez-Vilda, P. (2015). Towards the search of detection in speech-relevant features for stress. *Expert Systems*, 32(6), 710-718. <http://dx.doi.org/10.1111/essy.12109>
- Rodman, R. (1998). Speaker recognition of disguised voices: A program for research. *Proceedings of the Consortium on Speech Technology in Conjunction with the Conference on Speaker Recognition by Man and Machine: Directions for Forensic Applications* (pp. 1-22). Ankara, Turkey: COST250 Publishing Arm.
- Rose, P. (2002). *Forensic Speaker Identification*. London: Taylor y Francis.
- San Segundo, E. y Gómez-Vilda, P. (2014). Evaluating the forensic importance of glottal source features through the voice analysis of twins and non-twin siblings. *Language and Law/Linguagem e Direito*, 1(2), 22-41.
- Titze, I. R. (2000 [1994]). *Principles of voice production* (2nd edition). Iowa, City: National Center for Voice and Speech.
- Wolf, J. (1972). Efficient acoustic parameters for speaker recognition. *The Journal of the Acoustical Society of America*, 51(6B), 2044-2056. <http://dx.doi.org/10.1121/1.1913065>
- Zhang, C. y Tan, T. (2008). Voice disguise and automatic speaker recognition. *Forensic Science International*, 175(2), 118-122. <http://dx.doi.org/10.1016/j.forsciint.2007.05.019>