

**ANÁLISIS ACÚSTICO DE LA VOZ CREAKY DELIBERADA EN
MUJERES ESPAÑOLAS**

**ACOUSTIC ANALYSIS OF SPANISH FEMALE
DELIBERATE CREAKY VOICE PHONATION**

PILAR LIRIO

Universidad Nacional de Educación a Distancia (UNED)

pirlirio@gmail.com

Artículo recibido el día: 06/10/2015

Artículo aceptado definitivamente el día: 24/02/2016

Estudios de Fonética Experimental, ISSN 1575-5533, XXV, 2016, pp. 193-232

RESUMEN

Existen diferentes ajustes o mecanismos laríngeos para producir la voz *creaky*, cualidad de voz que combina un sonido similar a una crepitación laríngea con frecuencia fundamental. 40 hablantes españolas de entre 20 a 35 años de edad, variedad centropeninsular, graban muestras de voz *creaky* y voz modal sin haber recibido entrenamiento previo. Se utiliza el programa informático Praat para extraer las siguientes medidas acústicas de la voz modal de cada una de las hablantes y de su correspondiente voz *creaky*: la diferencia en hercios entre ambas F_0 y la diferencia en decibelios entre los dos primeros armónicos (H_1-H_2) y entre el primer armónico y el de mayor amplitud del tercer formante (H_1-A_3). Los resultados obtenidos revelan que la totalidad de las hablantes efectúan un descenso de la F_0 modal. En cambio, la diferencia entre H_1-H_2 para la voz *creaky* es positiva para un 28% y la diferencia entre H_1-A_3 resulta ser positiva para un 55% de las hablantes, alejándose ambos valores de los considerados prototípicos. Por último, se observan tres grupos diferenciados de ajustes laríngeos atendiendo al valor de descenso de F_0 y a la diferencia entre H_1-H_2 .

Palabras clave: *creaky*, *modal*, *frecuencia fundamental*, *armónicos*, *voz femenina*.

ABSTRACT

Creaky voice is a phonation type that can be produced by different laryngeal settings which are still under research. Forty young European Spanish female speakers with no previous training, recorded some samples in modal and creaky voice. Acoustic measures were extracted and analyzed using the Praat software in order to classify different phonatory strategies to produce a creaky voice. At first, we analyzed the F_0 values for modal and creaky voice and secondly, we focused on the amplitude difference between the first and the second harmonic (H_1-H_2) and between the first harmonic and the one with the highest amplitude of the third formant (H_1-A_3). The results revealed: the female speakers produced the creaky voice phonation by lowering their modal F_0 , 28% of the speakers showed a positive amplitude between H_1-H_2 and 55% of speakers had also a positive result between H_1-A_3 . These harmonics measures are contrary to the literature. Finally, considering the F_0 rate decreasing value and H_1-H_2 amplitude difference, we discovered three different groups of phonatory strategies that must be analyzed conducting further experiments.

Keywords: *creaky*, *modal*, *fundamental frequency*, *harmonics*, *female*, *Spanish*.

1. INTRODUCCIÓN

El habla es el principal medio de comunicación, por el cual los seres humanos transmiten información de muy distinta índole (Laver, 1994:21-23). Uno de los rasgos que transmite gran parte de la información es la cualidad de voz que son las características presentes en una voz durante todo el tiempo que la persona habla (Abercrombie, 1967:91; Laver, 1980) y que causan una determinada impresión perceptiva en el oyente (Kreiman y Sidtis, 2011; Hewlett y Beck, 2013:212). Así pues, la cualidad de voz es lo que caracteriza a un hablante (Gerratt y Kreiman, 2004).

La cualidad de voz procede de la configuración anatómica de la persona pero también puede originarse por la acción voluntaria ejercida sobre las diferentes estructuras laríngeas, sobre las estructuras supralaríngeas o sobre ambas estructuras en combinación (Gil, 2012). En el caso de tratarse de una acción voluntaria, se habla de cualidad de voz si dicha acción afecta continuamente a toda la forma de hablar de una persona (Laver, 1980; Trask, 1996:380).

La voz objeto del presente estudio, se describe perceptivamente de diversas formas: *el sonido de un palo al recorrer una verja* (Catford, 1964), *una sucesión de pulsos producidos por la laringe* (Hollien y Wendahl, 1968), un sonido *chirriante, crepitante*, como de *puerta oxidada, motor de un barco o palomitas de maíz explotando* (Moser, 1942; Henton y Bladon, 1988; Blomgren, Chen y Gilbert, 1998). Terminológicamente, plantea muchos problemas. Se le denomina *creaky voice*, *creak*, *vocal fry*, *gottal fry*, *glottalization*, *pulse register*, *laryngealisation*, *pulse phonation*, o *strobass*. No obstante, el fenómeno vocal es fácilmente identificable, ya que se pueden percibir los pulsos glotales de manera individual. Algunos autores defienden que todos los términos son sinónimos (Hollien, 1974), mientras que, para otros existen diferencias (Laver, 1980:160-161 y 1994:194-197 y Hewlett y Beck, 2013). Laver denomina voz *creaky* a la combinación de este tipo de cualidad de voz tan debatido con la producción simultánea de frecuencia fundamental (Laver, 1994:195)¹.

Laver sugiere, además, la posibilidad de una gran diversidad de ajustes laríngeos para producir voz *creaky* (Laver, 1994:195) entendiendo ajustes como la combinación de los mecanismos de longitud, aducción y tensión de los pliegues

¹ Para evitar, dentro de lo posible, los problemas terminológicos en este estudio se han elegido solo las muestras que presentaban frecuencia fundamental y se ha denominado voz *creaky* al tipo de voz analizada.

vocales (Frič, Šram y Švec, 2006). Redi y Shattuck-Hufnagel hablan de «glotalización» para referirse a la voz *creaky* y defienden que los hablantes recurren a diferentes ajustes para producirla y que se manifiestan en las diferentes características acústicas que presenta la onda glótica. Y no solo existen diferencias sustanciales interlocutores (debido a las diferencias de lengua, sexo o posición del segmento en la palabra o la frase) sino, lo más importante, existen importantes diferencias intralocutor. Es decir, el mismo hablante produce la voz *creaky* de diferentes maneras, lo que se infiere del análisis de las medidas acústicas extraídas de su voz (Redi y Shattuck-Hufnagel, 2001).

Un grupo de 40 mujeres jóvenes proporciona una serie de muestras primero con su voz habitual y, después, con voz *creaky*. Las hablantes solo reciben indicaciones y ejemplos de cómo producir voz *creaky* por parte del investigador. De esa manera, se obtienen un conjunto de muestras espontáneas de las que se espera encontrar diferencias acústicas relevantes en la onda glótica. Eso significaría que las hablantes españolas recurren a diferentes tipos de ajustes laríngeos para producir la voz *creaky*. ¿En qué consisten esos ajustes? ¿Cuántos tipos diferentes se pueden observar? ¿Es posible que exista algún ajuste que compartan la mayoría de las hablantes?

2. ESTADO DE LA CUESTIÓN

2.1. Problema terminológico: Calidad de voz, timbre, registros vocales y modos de fonación

De las cuatro propiedades del sonido, intensidad, tono (o frecuencia fundamental), timbre y duración, es el timbre, la calidad perceptiva, la más controvertida. Según la Acoustical Society of America², el timbre es aquella característica perceptiva que permite a un oyente juzgar si dos sonidos con la misma intensidad y frecuencia fundamental son distintos. Además añade que el timbre depende, primeramente, del espectro de frecuencias pero que también depende de la presión sonora o intensidad y de las características temporales del sonido o duración.

² American Standards Association & Acoustical Society of America. (1960): *American standard acoustical terminology: (including mechanical shock and vibration)*, American Standards Association.

El interés por el estudio de esas características perceptivas de la voz no es una cuestión reciente. A principios del siglo XX Drake menciona ya la confusión terminológica existente al utilizarse «timbre», «personalidad» o «cualidad de voz» de manera indistinta (Drake, 1937). En 1927, Sapir reconoce que, hasta ese momento, solo se tiene conocimiento de la cualidad de la voz o timbre desde un punto de vista perceptivo y que es necesaria una gran cantidad de estudios de tipo científico puesto que no se sabe qué mecanismos de producción hay detrás de las diferentes cualidades de la voz (Sapir, 1927).

La mayoría de autores consideran sinónimos los términos «cualidad de voz» y «timbre» y los utilizan indistintamente aunque, en general, se prefiere el término «cualidad de voz» (Laver, 1975; Dejonckere *et al*, 2001; Kreiman, Vanlancker-Sidtis y Gerrat, 2008; Hewlett y Beck, 2013:218). Sin embargo Crystal defiende la diferencia entre timbre y cualidad de voz, argumentando que el concepto timbre se aplica de una manera general a las resonancias de cualquier sonido y en sentido estricto, a las características inherentes de un sonido segmental (por ejemplo, timbre vocálico), mientras que «cualidad de voz» debe reservarse solo a las características inherentes según las cuales se producen contrastes lingüísticos (Crystal, 1976). Sundberg también habla de timbre, pero atribuye sus características a las dimensiones del tracto vocal y sostiene que los dos formantes³ más bajos del espectro influyen sobre la cualidad de las vocales mientras que la cualidad de la voz depende de los formantes más altos (Sundberg, 2013:61). Scherer puntualiza que el término «cualidad de voz» se refiere a lo que se percibe subjetivamente como timbre o color de la voz (Scherer *et al*, 2013).

En el sentido amplio de la acepción, se considera cualidad de voz al producto de varios factores. En primer lugar, el equipamiento vocal anatómico y fisiológico del hablante; en segundo lugar, los ajustes musculares voluntarios laríngeos, responsabilidad de la musculatura intrínseca de la laringe, que se encarga de regular la tensión longitudinal y la fuerza de aducción de los pliegues vocales (Hirose, 1971; Hewlett y Beck, 2013) y se suele denominar «modos de fonación» (Catford, 1964; Laver, 1968; 1994); y, en tercer lugar, los ajustes supralaríngeos o articulatorios, que son los efectuados por las estructuras supraglóticas: la faringe, la úvula, la lengua, la mandíbula y los labios (Laver, 1994:406-413). La forma en que se ensamblan y coordinan todas esas estructuras anatómicas laríngeas y

³ Formante: Zona de la escala de frecuencias en la que un sonido presenta una mayor concentración de energía. También puede definirse como cada una de las resonancias del conducto vocal (Gil, 1988).

supralaríngeas constituyen los diferentes tipos de ajustes que conforman las distintas cualidades de voz (Honikman, 1964; Laver, 1968, 1980; Esling, 1983).

En sentido estricto, se considera que la cualidad de voz se corresponde solo con los modos de fonación (Gobl, 1989; Kreiman y Sidtis, 2011:6). Scherer (1989) opina que la cualidad de voz está determinada tanto por los modos de fonación como por los registros fonatorios. *Cualidad fonatoria* es como algunos autores denominan la cualidad de voz (Gobl y Ní Chasaide, 1992; Keating y Esposito, 2006). Ladefoged habla indistintamente de cualidad de voz, modos de fonación y estados de la glotis y propone un modelo de cinco tipos de fonación en un continuo que va desde el cierre absoluto de los pliegues vocales hasta la apertura total (Ladefoged 1971:7; Gordon y Ladefoged, 2001).

En cuanto a los modos de fonación, Catford (1964:26) advierte del problema existente con el término «fonación» ya que los autores bien lo consideran como un fenómeno relativo a la producción del sonido, lo que implica la intervención del tracto vocal, o bien relativo a la producción de la voz. La confusión viene motivada porque, si bien la fonación es una actividad laríngea, si el sonido generado tiene valor fonológico, se considera articulatoria, en cambio si el sonido generado por la laringe necesita de la modulación supraglótica, la actividad es fonatoria. Catford clasifica los modos de fonación dependiendo del tipo y la localización del estrechamiento en la laringe. Los modos que dependen del tipo de estrechamiento son: *breath*, *whisper*, *voice*, *creak*, *stop* y los que dependen de la localización: *glottal*, *ligamental*, *arytenoidal*, *ventricular*. Aparte de estos modos, existen otros que resultan de la combinación del tipo y la localización del estrechamiento: *breathy voice*, *whispery voice*, *whispery creak*, *voiced creak* y *whispery voiced creak* (Catford, 1964:30-33).

Para Laver la fonación tiene que ver con la generación de energía acústica en la laringe debido a la acción de los pliegues vocales y divide los ajustes de la laringe en tres categorías: la frecuencia fundamental, la intensidad y los modos de fonación. (Laver, 1968; 1994:132). Laver también distingue entre modos de fonación simples: *modal voice*, *falsetto*, *whisper*, *creak*, *harshness* y *breathiness* y compuestos que se producen al concurrir dos o más modos de fonación simples como por ejemplo: *whispery voice*, *whispery falsetto*, *whispery creak*, *whispery creaky voice*, *whispery creaky falsetto*, *harsh whispery voice*, *breathy voice*, *creaky voice*... Además considera que existen combinaciones imposibles (Laver, 1980: 93-140).

El «registro vocal» describe un rango de frecuencias fundamentales consecutivas que se pueden producir con una cuasi homogénea cualidad de voz (Hollien, 1974; Titze, 1994:253). Hollien alterna el término «cualidad de voz» con el de «cualidad fonatoria» y propone tres registros básicos: voz pulsada, modal y falsete⁴. Titze, a la hora de enumerar las características del sonido, no habla de «timbre» sino de «cualidad de voz» y la considera constituida por todas aquellas características que se perciben de una voz al restarle los componentes de frecuencia fundamental e intensidad y eso incluye los «registros vocales»⁵ (Titze, 1994:252) y las modificaciones del tracto vocal (Story y Titze, 2002). Sveč (2000) realiza una revisión bibliográfica y clasifica los factores potenciales que pueden resultar importantes para la consecución de los registros vocales según los estudios que se centran exclusivamente en el análisis del mecanismo vibratorio de los pliegues vocales (como los de van den Berg, Hirano o Titze), los que analizan la resonancia del tracto vocal (como Large, Miller o Titze), los estudios dedicados a la interacción de las resonancias subglóticas y supraglóticas junto a la vibración de los pliegues por la influencia del filtro sobre la fuente (Titze o McAllister) y los que se centran en el estudio de los factores perceptivos (Colton, Titze o Sundberg).

En el presente trabajo se opta por el término «cualidad de voz», que es el aceptado por la mayoría de la bibliografía, y se considera cualidad de voz a la impresión psicológica creada por un estímulo físico y que depende de la voz y del oyente (Kreiman y Sidtis, 2011:6). La cualidad de voz se compone de un conjunto de ajustes glóticos y supraglóticos en el sentido de Laver y Honikman, los ajustes glóticos son los llamados «modos de fonación» y se considera fonación como la generación de energía acústica en la laringe debido a la acción de los pliegues vocales, según propone Laver. La clasificación de los modos de fonación elegida es la propuesta por Laver, en concreto, los modos de fonación que se estudian en el presente trabajo son la voz modal y la voz *creaky*.

⁴ Hollien denomina *loft* al registro de falsete (Hollien, 1974).

⁵ En la octava edición de la Conferencia sobre la Fisiología de los Pliegues Vocales (1994), Titze aporta una lista de elementos que integran la cualidad de voz entre los que se encuentran los relativos a los registros y en la página web del National Center for Voice and Speech aparece una tabla con los elementos integrantes de la cualidad de voz propuestos por Titze. También aparecen unas breves descripciones perceptivas y fisiológicas. Consultado en diciembre de 2015 en <http://www.ncvs.org/ncvs/tutorials/voiceprod/tutorial/quality.html>

2.2 Voz *creaky* y voz modal: descripción fisiológica

Los distintos ajustes laríngeos que dan lugar a los modos de fonación (Laver, 1968), son el resultado de las posibles combinaciones entre la aducción aritenoidea, la compresión medial y la tensión longitudinal de los pliegues vocales (figura 1).

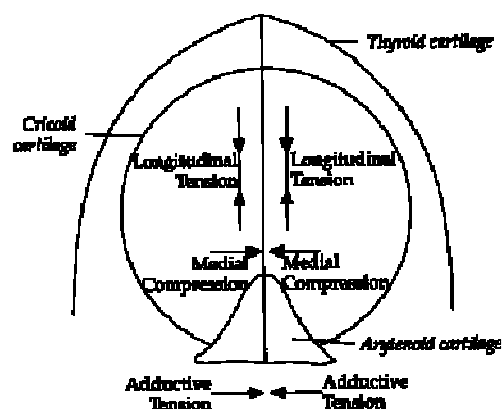


Figura 1. Ilustración que representa el corte sagital de la laringe. Las flechas indican la dirección de las tensiones musculares. (Gobl y Ní Chasaide, 2010:395).

Desde un punto de vista lingüístico, la configuración articulatoria para los diferentes modos de fonación no cuenta con la unanimidad de los investigadores. Eso es debido a que, como menciona Laver en sus publicaciones (1980, 1994) el hablante puede recurrir a ajustes diferentes para conseguir el objetivo fonatorio. La voz modal⁶ es, pues, un modo de fonación simple y consiste en la fonación normal o habitual de cada hablante, para la que se necesita un ajuste laríngeo basado en el

⁶ El término acuñado, en un principio, para este modo de fonación era el de voz normal, sin embargo, fue Hollien (1974) el que comenzó a utilizar el término voz modal debido a una puntualización efectuada de forma verbal por van den Berg en 1966 que sugirió que el término «normal» podía implicar la condición de no normalidad para el resto de modos de fonación. Laver habla más bien de normalidad en el sentido de posición neutra de los órganos vocales que permiten que los pliegues vocales vibren periódicamente, sin fricción audible y con una alta eficiencia (Laver, 1980:14).

alto grado de aducción aritenoides, compresión media moderada y mínima tensión longitudinal (Laver, 1980:111; Gobl y Ní Chasaide, 2010; Hewlett y Beck, 2013). El *creak* es también un modo de fonación simple cuyos ajustes laríngeos no están del todo claros. Parecen consistir en la combinación de un alto grado de aducción aritenoides y compresión medial, junto a un grado bajo de tensión longitudinal (Hewlett y Beck, 2013). Por último, la voz *creaky*, que es el objeto de investigación del presente estudio, resulta de la combinación simultánea de los dos modos simples explicados anteriormente, es decir, un modo de fonación en el que la voz *creak* se presenta con frecuencia fundamental (Catford, 1977:101, Laver, 1994:195; Hewlett y Beck, 2013:276).

La observación directa de los pliegues vocales no ha aclarado demasiado sobre el mecanismo para producir una voz percibida como *creaky*. Trask (1996) recomienda que se evite el uso indistinto de *creak* y *creaky* para referirse al mismo fenómeno. En el *creak* existe compresión lateral, solo vibra una pequeña parte anterior de los pliegues, mientras que las bandas ventriculares están aducidas e, incluso, presionan los pliegues vocales. En cambio, la voz *creaky* es un tipo de fonación compuesto en el cual parte de la glotis produce *creak* (ver figuras 3, 4 y 5) mientras otra parte produce voz modal. En este sentido, Teshigawara (2003) cita las investigaciones de Fujimoto y Maekawa (2003) que utilizan un fibroscopio para observar los estados de las áreas glótica y epiglótica, y observan que la aducción de las bandas ventriculares y la constricción de la cavidad faríngea producen un tipo de fonación comparable perceptivamente a la voz *creaky*.

Sin embargo, Esling y Harris (2005) y Edmonson y Esling (2006) opinan que este tipo de voz se debe principalmente a la acción de los repliegues ariepiglóticos unida a una vibración lenta de los pliegues vocales y la nula o casi nula intervención de los pliegues ventriculares. La configuración extremadamente estrecha en el esfínter ariepiglótico induce la vibración de dichos repliegues lo que, como en el caso anterior, también produce un tipo de fonación comparable perceptivamente a la voz *creaky*. Un ejemplo de ello se muestra en la figura 2.



Figura 2. Voz modal frente a voz creaky: imágenes endoscópicas. El mecanismo para producir voz creaky parece involucrar los pliegues ariepiglóticos (Gick, Wilson y Derrick, 2013:110).

2.3. Características acústicas de la voz *creaky*

El análisis acústico de una voz percibida como *creaky* produce mucho mejor resultado en la descripción de sus características que la observación directa ya que son muchas las medidas acústicas que se pueden extraer y los ajustes laríngeos a los que recurren los hablantes para producir voz *creaky* se pueden inferir del significado de dichas medidas acústicas.

Los dos principales sistemas de registro de la señal de voz para el posterior análisis acústico son: grabación de la señal radiada de los labios y registrada a través de micrófono y registro de la onda glótica a través de la señal electroglotográfica.

La señal electroglotográfica se obtiene colocando al paciente dos electrodos a ambos lados del cuello sobre la región donde se localizarían las láminas del cartílago tiroides. Cuando los pliegues vocales se aproximan para la fonación se produce un cambio de impedancia que queda registrado en forma de una onda. La técnica utilizada para la obtención de este tipo de señal se llama electroglotografía (EGG). (Fernández-Baillo, 2013).

Tradicionalmente se ha utilizado la electroglotografía (EGG en adelante) tanto para el estudio de las patologías vocales como para otros aspectos de la investigación sobre la voz ya que es un sistema de exploración indirecta. Si los tejidos están en contacto, hay conducción, si no lo están, hay aire y no hay conducción (Titze, 1990). Véase la figura 3.

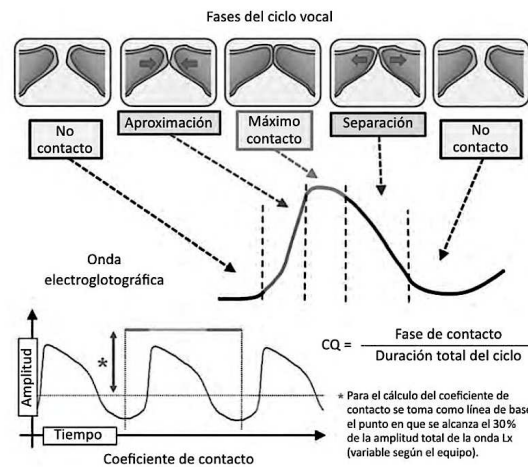


Figura 2
Interpretación del registro
electroglotográfico.

Figura 3. Onda electroglotográfica y su correspondencia con las fases del ciclo vocal. Adaptado de Cobeta et al. (2013:217).

2.3.1. Frecuencia fundamental

La frecuencia fundamental o tono responde a la vibración de la primera onda simple de las que forman la voz humana. La voz *creaky* se caracteriza, normalmente, por un descenso de la frecuencia fundamental modal. El descenso frecuencial de la voz *creaky* con respecto a la voz modal suele oscilar desde un cuarto de la frecuencia fundamental modal media de un hablante (Gick, Wilson y Derrick, 2013:109) hasta una octava por debajo de su frecuencia fundamental modal⁷ (Blomgren, Chen, Ng y Gilbert, 1998).

No obstante, aunque el descenso parece ser lo más común, al analizar las diferentes lenguas que usan voz *creaky* con valor fonológico, se han descubierto diferentes tipos de producción. El chino mandarín, por ejemplo, baja la frecuencia fundamental para producir voz *creaky* mientras que el bura hace justo lo contrario, eleva la frecuencia fundamental (Keating y Esposito, 2006).

Respecto al valor de F_0 necesario para catalogar una muestra como representativa de voz *creaky*, no hay acuerdo entre los autores. Catford es de los primeros en

⁷ De aquí en adelante se usarán indistintamente los términos F_0 y frecuencia fundamental.

establecer que el *creak* se encuentra por debajo de los 50 Hz. (Catford, 1977:101). Hewlett y Beck (2013:276) sitúan la frecuencia fundamental del *creak* en una zona que va de los 50 a los 90 Hz., de los valores más altos recogidos en la bibliografía y Gerrat y Kreiman (2001) y Kreiman y Sidtis (2011:276) contemplan un rango de F_0 de 7 a 78 Hz. para este modo de fonación, de los valores más bajos encontrados en los estudios. Entre estos valores, máximo de 90 Hz. y mínimo de 7 Hz. los autores han hallado una gran variedad de medidas. En general, y según lo que Michel (1968) sostiene, la frecuencia fundamental máxima de este modo de fonación al que él denomina *vocal fry* es de 100 Hz.

La frecuencia fundamental tiene un fuerte efecto sobre el modo de fonación aunque aún todavía queda por determinar si se trata de un fenómeno limitado a ciertas lenguas (Keating y Esposito, 2006), a ciertas vocales o si depende del sexo del hablante (Jiang *et al*, 2002). Aún no hay resultados concluyentes al respecto.

2.3.2. Armónicos

Los armónicos son las frecuencias que componen la onda compleja y que suponen los múltiplos de la fundamental (Gil, 1988). Los armónicos se pueden observar mediante un corte espectral en el que se muestran dos magnitudes, la amplitud del armónico y la frecuencia. La frecuencia fundamental es el armónico primero o H_1 , el que tiene mayor amplitud. A medida que los armónicos evolucionan en frecuencia, disminuyen en amplitud. Este fenómeno se denomina declinación espectral (véase la figura 4).

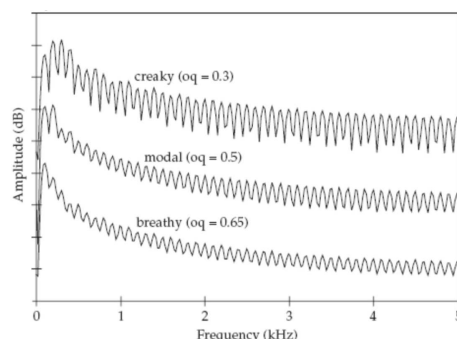


Figura 4. Análisis espectral de tres modos de fonación. En la imagen se puede observar la declinación espectral. En la voz creaky, se observa que

el primer armónico tiene menor amplitud que los siguientes armónicos que aparecen reforzados. Adaptado de Johnson (2012:172).

Sin embargo, una de las características de la estructura armónica de la voz *creaky* consiste en el refuerzo de la amplitud de los armónicos más altos, lo que cambia el patrón de la declinación espectral. En este caso, el segundo armónico (H_2) tiene mayor amplitud que el primero (H_1) (Keating y Esposito, 2006; Kreiman y Sidtis, 2011; Johnson, 2012). Holmberg vinculó la fuerte relación entre los armónicos H_1 - H_2 con las fases de apertura y cierre de la fuente glótica (Holmberg *et al*, 1995). H_1 tiene como correlato glótico la apertura de los pliegues vocales, a mayor cociente de apertura dentro de un ciclo glótico, mayor amplitud del H_1 (Kreiman *et al*, 2012). En la voz modal, la fase de cierre y la fase de apertura de los pliegues dentro un ciclo glótico, es equivalente en cuanto a duración, por eso la relación entre H_1 - H_2 es de igualdad en amplitud (Blomgren, Chen, Ng y Gilbert, 1998; Johnson, 2012:170). En el caso de la voz *creaky*, el coeficiente de cierre es superior al de apertura, por eso, el segundo armónico es mucho más amplio que el primero. Este hecho es el que causa que se perciban los pulsos glóticos de manera individual. Un coeficiente de apertura corto produce el descenso de la frecuencia fundamental o H_1 , mientras que las frecuencias agudas del espectro se ven fortalecidas (Holmberg *et al*, 1995; Gordon, 2001; Keating y Esposito, 2006; Erickson *et al*, 2008). En este sentido, los valores muy bajos en la diferencia entre H_1 - H_2 se correlacionan con un incremento en la constricción de la glotis (Seyfahrt y Garellek, 2015).

La diferencia entre la amplitud del primer armónico y el armónico de mayor amplitud del tercer formante, A_3 (H_1 - A_3) también demuestra el refuerzo de los armónicos más altos (Hanson y Chuang, 1999; Iseli *et al*, 2007) y su correlato glótico es el cociente de velocidad o *skewness* que refleja la asimetría del pulso glótico (Hanson, 1995) y que consiste en la relación entre la velocidad en que los pliegues vocales tardan en cerrarse y en abrirse. En concreto, en la voz *creaky*, el cierre de los pliegues resulta muy brusco y la velocidad de cierre presenta índices muy elevados, mientras que la fase de apertura es muy lenta (Stevens, 1977, citado por Esposito, 2010; Luo, 2012). Véase la figura 5.

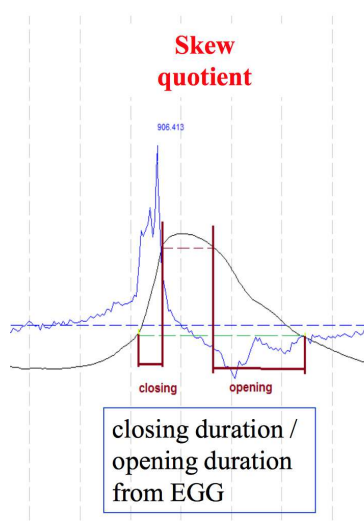


Figura 5. Imagen donde se puede observar la onda EGG en color negro. El cociente de velocidad o skew que es el cociente entre las velocidades de cierre y apertura del pliegue. Se observan dos fenómenos, el primero consiste en que la fase de cierre es menor que la fase de apertura. Acústicamente, el parámetro que lo evidencia es la diferencia negativa entre H_1-H_2 . Además, si se suman la fase de cierre, la de apertura y la parte central que es el tiempo en que el pliegue está cerrado, se advierte que la fase total de cierre del pliegue es superior a la fase de apertura. El segundo fenómeno es que la fase de apertura, al ser mayor en duración que la de cierre, significa que es mucho más lenta, lo que se refleja en la diferencia también negativa entre H_1-A_3 (Adaptado de Keating et al, 2012).

La velocidad en el cierre de los pliegues es el factor que produce el aumento de amplitud de los armónicos que se encuentran en la región del tercer formante (Morris y Harmon, 2010). Acústicamente, si la diferencia entre H_1-A_3 es negativa, es decir, los armónicos más altos son superiores al primero y la estructura armónica no decae de forma regular como ocurre en la voz modal, entonces se produce voz *creaky* (Erickson et al, 2008). Precisamente, al estar relacionada la velocidad de cierre con el refuerzo de las frecuencias altas del espectro, se

considera una medida alternativa para distinguir algunos modos de fonación, cuando H_1-H_2 falla (Keating y Esposito, 2006).

En un estudio sobre los diferentes modos de fonación en Santa Ana del Valle Zapotec (Esposito, 2010), la diferencia entre H_1-H_2 fue más relevante en hablantes femeninas a la hora de distinguir entre voz *creaky* y modal que la diferencia entre H_1-A_3 al contrario de lo que ocurre en voces masculinas. Quizás esto sea debido bien al hiato presente entre los pliegues vocales de las mujeres que impide que exista un cierre glótico completo o bien a que la configuración laríngea usada por uno y otro sexo da lugar a esta disparidad (Esposito, 2010).

2.4. Tipos de voz *creaky*

Desde un punto de vista lingüístico y acústico (no así fisiológico), se observa que la mayoría de hablantes que producen voz *creaky* comparten una estrategia fonatoria consistente en los ajustes expuestos en los anteriores apartados⁸ que pueden resumirse (con respecto a los parámetros analizados en el presente estudio) del siguiente modo: descenso considerable de la F_0 y espectro armónico reforzado en las frecuencias altas que se determina por la relación negativa entre H_1-H_2 y H_1-A_3 producto de un alto grado de constricción laríngea y que tiene su correlato glótico en la lenta y abrupta fase de cierre de los pliegues vocales en comparación con la de apertura. Esta configuración laríngea es denominada voz *creaky* «prototípica» por Keating, Garellek y Kreiman (2015) que proponen el término para distinguir esta configuración de otras variedades observadas por los investigadores.

Hedelin y Huber en 1990 diferencian entre: *creak* o *fry* (con F_0 bajo), voz *creaky* (con pulsos irregulares) y diplofonía (con doble período). Redi y Shattuck-Hufnagel en 2001 distinguen hasta cuatro tipos de voz *creaky*: irregular aperiódica, amortiguada, diplofónica y la inusual voz *squeak* (con F_0 elevado). También se ha observado que cuando F_0 es elevado, como suele ocurrir en posición inicial o aislada, la voz *creaky* pierde su característico sonido pulsátil (Esposito, 2010).

Keating, Garellek y Kreiman en 2015, proponen la siguiente clasificación: *vocal fry* en el que la glotis se encuentra constreñida, el F_0 es bajo aunque la vibración no es necesariamente irregular y la fase de cierre de los pliegues vocales es más larga que en el resto de tipos; la voz pulsada múltiple que produce ciclos largos y breves

⁸ Véanse los apartados 2.2 y 2.3.

que se alternan, no se observa una F_0 especialmente baja y los pulsos tienen una fase de cierre bastante larga; la voz tensa o presionada en la que el F_0 es regular y no demasiado bajo pero el grado de constricción glótica es muy alto; la voz aperiódica en la que no se percibe frecuencia fundamental, no hay periodicidad y la constricción glótica es alta; y el *creak* no constreñido o «Slifka» (denominado así por ser esta investigadora la primera en describir este tipo), en el que la F_0 es baja e irregular pero no hay constricción y la glotis, incluso, deja salir un escape de aire. En la figura 6 se puede observar un esquema de su planteamiento.

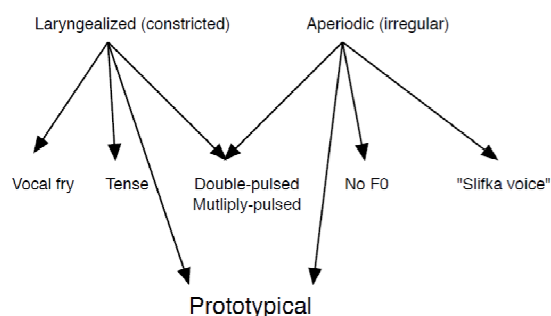


Figura 6. Tipos de voz creaky según Keating y Garellek (2015). Adaptado del póster para el congreso anual de la Liguistic Society of America, 2015.

La clasificación adoptada por Keating, Garellek y Kreiman (2015) es la elegida para catalogar las muestras analizadas en el presente estudio por ser la más actual. Sin embargo, es necesario señalar que el presente estudio solo analiza la existencia o no de F_0 y el análisis de la diferencia entre H_1 - A_3 .

2.5. Funciones de la voz creaky

La voz *creaky* viene determinada, en muchas ocasiones, por las diferencias en cuanto a la conformación física del hablante, su edad, sexo o clase social (Beck, 2010). Tiene diferentes usos, desde fonológicos y emocionales hasta socio-lingüísticos. Se utiliza en la voz cantada, como indicador en fonética forense y es muy común en patologías vocales.

La voz *creaky* suele considerarse como revelador de la existencia de una posible patología vocal (Hanson, Stevens, Kuo, Chem y Slifka, 2001). Las patologías vocales afectan a las características vibratorias de los pliegues vocales manifestándose en las medidas aerodinámicas y acústicas de la voz. Y también muestran ciertas características en el ámbito perceptivo. Por eso, la voz *creaky* se usa como uno de los parámetros a estimar en algunas escalas de evaluación vocal perceptiva como la *Stockholm Voice Evaluation Approach* (SVEA) (Hammarberg y Gauffin, 1995; Hammarberg, 1999) o el *Vocal Profile Analysis Scheme* (VPA) (Laver, Wirz, Mackenzie y Hiller, 1981). Sin embargo, la valoración de la existencia de una patología vocal no depende de la presencia o no de la voz *creaky* sino de su grado y de la concurrencia de muchos más factores (Fawcus, 1986).

Desde el punto de vista de la fonética forense, el disimulo vocal es un problema importante a la hora de identificar al hablante. El 52% de delincuentes optan por alguna forma de enmascaramiento de sus rasgos fónicos. Dentro de los métodos no electrónicos de disimulo, se encuentran las alteraciones de los modos de fonación, sobre todo, la frecuencia fundamental en general y el uso de la voz *creaky* en particular (Rodman, 1998; Perrot, Aversano y Chollet, 2007). En el caso de la voz femenina, Moosmüller (2007) estudia el efecto de este tipo de enmascaramiento en las frecuencias formánticas y descubre que existen diferencias en cuanto al sexo del hablante: la voz *creaky* producida por voces femeninas muestra un segundo formante más bajo comparado con el mismo segmento producido usando fonación modal. El reto es, por tanto, encontrar indicadores para detectar el tipo de disimulo y descubrir qué características de la onda glótica permanecen inalterables (Perrot, Aversano y Chollet, 2007).

Una de las principales funciones de la voz *creaky* es la producción de contrastes fonológicos, por ejemplo, para diferenciar vocales en lenguas amerindias (Ladefoged, 1971:15; Belotel-Grenié y Grenié, 2004) o como ocurre con el chino mandarín y el bura (una lengua nigeriana) (Keating y Esposito, 2006).

Sin embargo, en gran cantidad de las lenguas del mundo, es importantísima su función paralingüística como expresión de emociones y evocación de sentimientos. Johnstone y Scherer (1999) estudian a través de datos obtenidos por electroglotografía, cambios en la onda glótica debido a diferentes estados emocionales. La voz *creaky* parece asociarse con aburrimiento en caso de hablantes de lengua inglesa como señala Laver (1980) citado por Gobl y Ní Chasaide (2003) y por Yanushevskaya, Gobl y Ní Chasaide (2005). Cullen, Kane, Drugman, y Harte (2013) encuentran una fuerte correlación entre la voz *creaky* y

algunos marcadores afectivos de una batería de estímulos extraída de la base de datos SEMAINE⁹.

Sociolingüísticamente, se conoce hoy que las mujeres norteamericanas emplean la voz *creaky* más que los hombres sin importar su edad o su raza como observa Podesva (2011). Algunos autores explican el origen de este uso como una táctica para parecer más autoritarias (Tajfel, 1974; Coates, 1986; Dilley *et al.*, 1996; Yuasa, 2010). La voz *creaky* aparece, incluso, en el inglés chicano (Fought, 2010; Mendoza-Denton, 2011) y en el español hablado por los hablantes de herencia¹⁰ de Estados Unidos (Kim, 2013). E incluso, el uso de la voz *creaky* por mujeres angloparlantes parece tener rasgos persuasivos o seductores si se usa con un interlocutor masculino (Shaw y Crocker, 2015). Por otro lado, un estudio sobre cualidad de voz y sexo en población hablante de portugués brasileño señala que prevalece el uso de la voz *creaky* entre los hombres (Camargo, Madureira, Pessoa y Rusilo, 2012).

Por último, en el campo de la voz cantada, se considera que los cantantes deben entrenar este modo de fonación para emplearlo de manera cómoda y saludable desde un punto de vista artístico (Henrich, 2006; Miller, 2004). Se prefiere utilizar los términos *fry* o *strobass* para referirse a este fenómeno.

3. METODOLOGÍA

3.1. Estímulos y sujetos

El corpus sobre el que se ha trabajado es el recogido en el proyecto CIVIL (Cualidad Individual de la Voz en la Identificación de Locutores) del Laboratorio de Fonética del CCHS¹¹. Este corpus se compone de grabaciones de 50 mujeres y

⁹ En concreto, se analizan los estímulos de valencia, activación, dominancia y expectación.

¹⁰ Hablantes de herencia (*heritage speakers*) son aquellas personas expuestas a una lengua que se habla en su casa y que no es la mayoritaria de la sociedad donde se están educando (adaptado de <http://nhlrc.ucla.edu/nhlrc/page/about>, consultado el 14 de julio de 2015)

¹¹ CCHS son las siglas que corresponden al Centro de Ciencias Humanas y Sociales perteneciente al CSIC. La página web informativa del proyecto CIVIL es: <http://www.estudiosfonicos.cchs.csic.es/fonetica/civil.php>

50 hombres centro-peninsulares de entre 20 y 35, cuya lengua materna es el español europeo estándar (San Segundo, Alves, y Trinidad, 2013). La muestra elegida para el presente análisis corresponde exclusivamente a voces femeninas.

La razón de dicha selección es que, al no existir acuerdo entre los autores sobre el límite en hercios que diferencian el *creak* y la voz *creaky* y como el objeto del trabajo es el estudio acústico de la voz *creaky* en comparación con la modal, es imprescindible asegurar que las voces analizadas no carezcan de F_0 para el correcto cálculo del sistema armónico. Por tanto, la voz femenina presenta en producción modal una frecuencia fundamental mucho más elevada que la masculina, así pues, se presupone que el descenso del F_0 no alcanza los límites conflictivos. Por el mismo motivo, han sido descartadas las muestras que carecen de F_0 o en las que el algoritmo del programa informático se ha mostrado ineficaz para calcular alguno de los parámetros acústicos objeto del estudio.

El corpus está constituido por dos grabaciones de cada informante con un mes de diferencia entre ambas muestras. Para este trabajo se ha elegido trabajar con la primera grabación por considerarse que no ha existido posibilidad de entrenamiento por parte de las informantes y lo que se estudia son los ajustes laríngeos *naturales* a los que recurren las hablantes no entrenadas. Las informantes no muestran evidencias de problemas de audición ni de voz.

Una de las tareas que los informantes deben llevar a cabo es la lectura de 33 palabras (de ellas, algunas son pseudopalabras) dentro de una frase marco, primeramente con voz modal y después con diferentes modos de fonación uno de los cuales es la voz *creaky*. Las palabras tienen estructura 'CV.CV.CV y la frase marco es: «Diga CV.CV.CV despacio varias veces». C es una consonante sonora no nasal para evitar nasalización de la vocal adyacente y V es la vocal [a]. La vocal elegida se halla en posición de sílaba tónica, libre y sin coda, precedida de la aproximante, velar, sonora [uj] en la palabra «gálibo». Se ha evitado que la vocal analizada se encuentre en posición de final de palabra, porque eso podría afectar a la cualidad de voz. Además, se elige la vocal [a], porque en esta vocal, H_1 no se encuentra apenas influido por la acción de las cavidades supraglóticas. Así pues, una vez seleccionadas las muestras, el análisis supone realizar un corte espectral en la mitad del segmento en la vocal con producción modal y en la vocal con producción *creaky*. Debido a la escasa duración de los segmentos estudiados, resulta desaconsejable el análisis perceptivo de dicho segmento.

De las 50 grabaciones disponibles, se ha descubierto un fallo de etiquetación siendo realmente 49 las informantes que conforman el corpus. Dos de las

informantes no llegan a decir la palabra «gálibo» en las muestras de voz *creaky* por lo que se rechazan para el análisis. Así pues, son 47 muestras de habla femenina las que pasan a la fase de análisis.

3.2. Recogida de datos

3.2.1. Equipo de grabación

Los datos se han recogido pidiendo a las informantes que leyeran de la forma más natural posible los estímulos contruados. Se les pidió efectuar una serie de tareas de producción relacionadas con la alteración voluntaria de los modos de fonación. Para ello, el entrevistador les puso el ejemplo y las informantes trataron de realizar el modo de fonación solicitado. Los estímulos fueron presentados de manera aleatoria y se pidió a las participantes una sola lectura en modo de fonación modal y en modo de fonación *creaky*. Las muestras analizadas corresponden, pues, a la impresión perceptiva de la voz *creaky* de las informantes.

Los estímulos fueron grabados en la cabina de grabación del Laboratorio de Fonética del CCHS, mediante un micrófono de condensador E6i Omnidireccional Earset Audio y una tarjeta de sonido UA-25EX de Roland. La señal fue grabada directamente en un ordenador de mesa con sistema operativo Windows en formato .wav con el programa Adobe Audition 1.0, ajustado a una frecuencia de muestro de 44.100 Hz. y a una resolución de 16 bits, mono.

3.2.2. Descripción de los procedimientos e instrumentos utilizados y de los criterios de análisis empleados

El programa Praat (Boersma y Weenick, 2013) utiliza un sistema de filtrado inverso de la señal para el cálculo de la frecuencia fundamental. Esta es la herramienta informática elegida para el análisis acústico. El filtrado inverso, mediante un algoritmo, calcula la derivada del pulso glótico para luego, volver a integrar el pulso glótico y, así, anular el efecto de radiación labial (Javkin, Antoñanzas-Barroso y Maddieson, 1987). Se utiliza esta herramienta porque en el presente trabajo se analizan parámetros de la acción glótica. Praat se utiliza, en primer lugar, para la extracción en formato .wav de los segmentos objeto de análisis. El segundo paso es la detección y cálculo de F_0 . La frecuencia fundamental se ajusta según el rango estándar de 75 Hz. a 500 Hz.

Por último, se extraen las diferencias entre los armónicos H_1-H_2 y H_1-A_3 y se utilizan para el estudio las medidas corregidas al ser las medidas más precisas para el análisis de la fuente glótica. Las medidas de los armónicos son muy sensibles a la altura frecuencial de los formantes. Las medidas corregidas, sirven para eliminar el efecto de la radiación labial y la influencia de los formantes, normalizando los datos de análisis ya que los exime de la influencia acústica del sistema articulador (Gobl y Ni Chasaide, 2010; Iseli *et al*, 2007).

Asimismo, se ha excluido del análisis la diferencia entre H_1-A_2 ya que el programa Praat no proporciona medidas corregidas para este parámetro y, si bien es cierto que el estudio sobre una vocal como la [a] asegura que el segmento esté libre de la acción de las cavidades supraglóticas, se considera recomendable desde el punto de vista de la precisión del experimento, no mezclar medidas corregidas con medidas no corregidas.

No obstante, la extracción de las medidas glóticas (frecuencia fundamental y armónicos), a través del filtrado inverso de señal, en ocasiones presenta fallos de detección relacionados con la irregularidad de la onda glótica. Esto es debido a que los algoritmos de detección de la frecuencia fundamental de los programas informáticos funcionan muy bien en el caso de que las vibraciones presenten regularidad, pero no en el caso de irregularidad como ocurre con la voz *creaky*. Se puede entonces deducir que si la detección de la frecuencia fundamental no es óptima, los armónicos no serán medidas robustas en las que se pueda confiar. (Cullen *et al*, 2013). Si afecta a la frecuencia fundamental, la informante ha de ser excluida del análisis, puesto que de la frecuencia fundamental dependen los armónicos por ser sus múltiplos. Si el fallo implica solo a algún armónico, se han realizado los análisis posibles. En este sentido, en la fase de extracción de la frecuencia fundamental se detectan algunos datos atípicos:

1. Informantes M07S01, M13S01, M15S01, M32S01, M41S01 y M45S01. El programa Praat no detecta frecuencia fundamental en el segmento objeto de análisis. No obstante, se procede a un examen visual del oscilograma con el programa Adobe Audition CS6 en el que se observa que solo la informante M15S01 no presenta ningún tipo de onda periódica, es decir, no presenta frecuencia fundamental. Para el resto de informantes, a pesar de que los cálculos efectuados sobre una de las ondas periódicas de la parte central del segmento objeto de análisis dan como resultado una frecuencia fundamental muy grave para las informantes M07S01 y M13S01, como el programa Praat, que es el elegido para la

realización de los análisis, no detecta frecuencia fundamental, se procede a eliminar los datos de estas informantes para el presente estudio.

2. La informante M19S01 presenta una frecuencia fundamental cercana al doble de su F_0 modal. Posiblemente se deba a fallos del programa, ya que los algoritmos de detección de la frecuencia fundamental del programa se revisan y mejoran continuamente. Al observar el oscilograma se observa una frecuencia fundamental para la voz *creaky* cercana a la mitad de su frecuencia fundamental en voz modal. Consultados expertos en detección de señales sobre este punto, la conclusión es que el fallo detectado en el algoritmo del programa Praat puede ser debido a la presencia de falsos cierres durante la fonación. En cambio, al usar Matlab como herramienta de detección de la frecuencia fundamental, se confirman los valores de F_0 que aparecen en el análisis oscilográfico. Se decide excluir este caso al ser imposible la detección de la frecuencia fundamental utilizando el programa Praat.
3. Las informantes M08S01 y M14S01 presentan un A_3 para voz *creaky* no definido, por lo que solo se analiza la diferencia H_1-H_2 .

Por ello, finalmente el análisis se ha basado en 40 muestras de voz *creaky* femenina, comparada con su correspondiente producción modal de las cuales, en dos de ellas solo ha sido posible realizar el análisis de la diferencia H_1-H_2 .

Para el tratamiento estadístico de los datos se utiliza el programa de IBM SPSS (2008). De los análisis estadísticos se excluyen los casos en los no se detecta frecuencia fundamental. La extracción del resto de los datos se realiza con el mismo programa informático.

4. ANÁLISIS E INTERPRETACIÓN DE LOS DATOS OBTENIDOS

En lo que respecta a la normalidad de la muestras, los resultados del test de Shapiro-Wilk muestran una distribución normal en todas las variables analizadas, excepto para la variable de diferencia entre H_1-H_2 para voz *creaky*. El coeficiente de asimetría para esta variable es -0,8 y el valor del índice de asimetría resultante de dividir el índice de asimetría entre su error estándar es de 2,3 (error estándar 0,3). El índice de simetría para el caso de menos de 200 muestras debería ser 2,58.

Por otro lado, el valor del índice de curtosis es de -0,70 (error estándar 0,7), no supera el índice 1,96 (Field, 2013). De los resultados anteriores unidos a la aplicación del teorema del límite central que garantiza una distribución normal cuando el tamaño muestral supera las 30 muestras, se infiere un grado moderado de asimetría y se asume la normalidad de la variable.

La prueba t de Student para muestras relacionadas entre la frecuencia fundamental modal y la de la voz *creaky* evidencia diferencias significativas entre ambas variables ($t(39) = -13,86$ $p = ,0001 < ,05$).

Como se ha explicado en el marco teórico, una de las principales estrategias fonatorias para producir voz *creaky* es el descenso en la frecuencia fundamental modal del hablante, lo que implica la relajación de la tensión longitudinal de los pliegues vocales. Las muestras de las informantes presentan descenso de F_0 de la voz *creaky* producida con respecto a su voz modal, a excepción de una de ellas que no presenta ningún tipo de descenso lo que puede significar un fallo de detección del programa o un dato atípico. El rango de descenso va desde los cero hercios, es decir, ningún tipo de descenso, hasta los -163 Hz. La media de la frecuencia fundamental de las voces *creaky* es de 129,8 Hz. con una desviación típica de 38,1, mientras que la media para la voz modal es de 220,8 Hz. con una desviación típica de 24,4. La media en hercios del descenso de F_0 es de -91 Hz. con una desviación típica de 41,5. La media en porcentaje es del 40%, lo que implica un porcentaje de descenso muy alto.

Respecto a si ambas variables se correlacionan, es decir, si cuanto más baja es la frecuencia fundamental de la voz modal, más baja es la de la voz *creaky*, existe una correlación muy baja entre las variables paramétricas F_0 *creaky* y F_0 modal, que no es significativa ($r = ,178$ y $p = ,272$). Esto significa que aunque las hablantes descienden la frecuencia fundamental modal para producir voz *creaky*, la frecuencia en sí de dicha voz *creaky* no depende de la frecuencia de la voz modal. Se pueden obtener valores muy bajos de voz *creaky* con valores altos de frecuencia fundamental modal.

En el caso del análisis de los armónicos, el análisis de frecuencias muestra que la media para la diferencia de H_1-H_2 para la voz *creaky* es de -4,66 dB con una desviación típica de 6,9. El ajuste laríngeo en este caso implica que la fase de apertura de los pliegues vocales es inferior a la fase de cierre. Sin embargo, analizando las muestras una a una se observa que 28 casos presentan una diferencia negativa, como se esperaba, un caso no presenta diferencia y 11 casos presentan diferencia positiva. Esto significa que el 27,5% de las hablantes

analizadas efectúa una serie de ajustes laríngeos para producir voz *creaky* diferente a la prototípica consistente en que la fase de apertura de los pliegues vocales es superior a la fase de cierre. Cuatro de estos 11 casos presentan, además, diferencia negativa para H_1-A_3 , es decir, que muestran refuerzo en la zona alta del espectro armónico, rasgo característico del tipo de fonación analizado. Y ya que la bibliografía (Keating y Esposito, 2006) lo considera una medida alternativa para cuando H_1-H_2 falla, se puede decir que solo 7 casos producen un tipo de fonación diferente (un 17,5%). En cuanto a la muestra que no presenta diferencias entre H_1-H_2 su valor H_1-A_3 es negativo, así pues, se puede inferir que la voz resulta acústicamente *creaky* ya que el espectro se ve reforzado en su zona alta y que la falta de diferencia entre H_1-H_2 puede significar un dato erróneo por el algoritmo del programa o un dato atípico. Véase la figura 7.

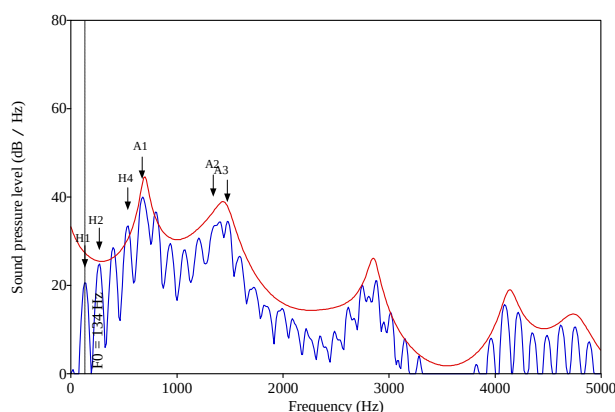


Figura 7. Corte espectral de la informante M26S01 en voz *creaky*, donde se puede observar gráficamente la diferencia entre los armónicos y la declinación espectral.

La media para la diferencia de H_1-H_2 en voz modal es, en cambio, un valor positivo, como se esperaba en todos los casos: 2,94 dB con una desviación típica de 2,1. Solo una de las hablantes presenta una diferencia negativa en voz modal y también la presenta para el valor H_1-A_3 . Puede existir la duda de si esta hablante ya utiliza un mecanismo de voz *creaky* en vez de modal para su voz habitual puesto que su voz *creaky* presenta valores muy negativos de H_1-H_2 y de H_1-A_3 . Además se trata de una de las hablantes con mayor diferencia de descenso de frecuencia

fundamental en hercios entre su voz modal y *creaky*. La frecuencia fundamental de su voz modal es de las más elevadas y la de su voz *creaky* es de las más bajas. Se procede a un análisis perceptivo de la hablante por parte de una logopeda (profesional de más de 30 años de experiencia) especialista en rehabilitación y patologías vocales. La especialista corrobora que se trata de una voz completamente sana sin rastro alguno de voz *creaky*. Entra dentro de lo posible que sea un fallo relacionado con el algoritmo usado por el programa informático.

En cuanto a la media de descenso para la diferencia H_1-A_3 en voz *creaky*, resulta ser de 0,93 dB con una desviación típica de 15,5, alejándose de lo considerado prototípico. Analizando las frecuencias por separado, 18 informantes muestran una diferencia negativa, 20 una diferencia positiva y en dos hablantes, el programa informático fue incapaz de calcular la diferencia de H_1-A_3 . Eso supone que el 50% del total de las muestras válidas efectúa una serie de ajustes laríngeos que suponen una diferencia positiva de H_1-A_3 en vez de la negativa esperable. H_1-A_3 indica la velocidad de cierre de los pliegues siendo el prototipo cierre brusco y rápido y apertura lenta, así pues, ese 50% de informantes actúa de manera contraria. Si se observa el parámetro H_1-H_2 para estos 20 casos, en 13 de ellos la diferencia es negativa y en 7 casos es positiva (17,5%). Este parámetro mide el cociente de apertura, así pues en 13 casos hay un mayor cociente de cierre y en 7 casos se produce un modo de fonación que parece tener algo de escape de aire puesto que el cociente de apertura es alto. Sin embargo, al escuchar las muestras completas, no se observa escape de aire. Puede ser que sea debido a que el análisis es del corte espectral en el medio de la vocal analizada.

La media de la diferencia de H_1-A_3 para la voz modal es de 3,66 dB con una desviación típica de 12,7 y se encuentran 17 muestras (42,5%) con diferencia negativa lo que también se aleja de las medidas prototípicas para voz modal. De la misma manera que en el caso anterior, con un vistazo al parámetro H_1-H_2 se observa que solo uno de estos 17 casos presenta valores negativos. Se trata de la hablante de la que anteriormente se dijo que posiblemente su voz hablada modal fuera *creaky*.

La tabla 1 refleja los datos extraídos de las informantes y está ordenada según el porcentaje de descenso de F_0 de menor a mayor.

Informante	F0 modal	F0 creaky	Diferencia F0	%Descenso	H1-H2	H1-A3
1	211	211	0	0,0	1,1	-1,0
32	220	191	-29	-13,2	0,6	22,6
17	270	234	-36	-13,3	-2,8	-1,1
5	178	153	-25	-14,0	-2,2	-13,4
10	225	184	-41	-18,2	-3,0	-5,8
30	196	157	-39	-19,9	1,4	20,4
15	204	162	-42	-20,6	-0,9	8,3
6	195	153	-42	-21,5	0,0	-9,1
31	185	139	-46	-24,9	-0,6	9,3
40	235	175	-60	-25,5	-0,5	15,1
16	230	169	-61	-26,5	2,6	-17,0
19	180	129	-51	-28,3	0,5	15,6
3	190	132	-58	-30,5	-1,7	-17,9
20	229	158	-71	-31,0	5,6	21,5
35	223	153	-70	-31,4	1,4	21,6
12	248	155	-93	-37,5	-2,7	-9,8
14	239	147	-92	-38,5	-0,8	5,0
4	237	143	-94	-39,7	0,8	-10,0
24	215	129	-86	-40,0	-1,1	-11,6
21	238	135	-103	-43,3	-3,0	-18,5
13	225	118	-107	-47,6	-2,2	-4,5
18	189	99	-90	-47,6	-12,3	-32,1
8	212	107	-105	-49,5	-11,5	-25,4
9	212	105	-107	-50,5	-9,9	-19,7
39	273	132	-141	-51,6	-2,1	14,4
37	238	115	-123	-51,7	-15,7	10,7
28	209	99	-110	-52,6	-5,2	18,4
22	203	96	-107	-52,7	-2,1	-9,8
26	240	113	-127	-52,9	-3,8	9,7
25	265	120	-145	-54,7	1,6	14,1

11	218	97	-121	-55,5	-16,6	
38	226	99	-127	-56,2	-12,5	-6,3
27	209	90	-119	-56,9	-10,7	5,3
23	260	109	-151	-58,1	-1,2	6,0
34	205	84	-121	-59,0	1,6	33,3
29	194	78	-116	-59,8	-14,1	1,0
33	250	96	-154	-61,6	-20,4	12,6
36	198	74	-124	-62,6	-15,3	3,6
2	218	75	-143	-65,6	-14,0	-20,0
7	240	77	-163	-67,9	-19,4	

Tabla 1. *Datos de todas las informantes ordenados de menor a mayor según el porcentaje de descenso de la frecuencia fundamental.*

Analizando los porcentajes de descenso en relación con las demás medidas acústicas se observa que 22 informantes presentan un porcentaje de descenso inferior al 50% de su voz modal y 17 presentan el 50% o más. Sin embargo, de estos 17 informantes, 13 se distribuyen en porcentajes que van del 50,5% al 59,8%. Es decir, en proporción, la mayoría de las hablantes bajan en torno al 50% de su voz modal.

También se observa que los valores con porcentaje bajo de descenso y hasta el 40% de descenso, 9 hablantes presentan medidas positivas y 9 medidas negativas de H_1-H_2 . Sin embargo a partir del 40% de descenso, solo se encuentran 2 valores positivos frente a 20 negativos. Es decir, cuanto mayor sea la diferencia entre ambas frecuencias fundamentales, mayor cantidad de valores negativos se dan en H_1-H_2 .

De hecho, existe una correlación positiva, moderada y significativa entre la variable de diferencia de ambas frecuencias fundamentales y la diferencia entre H_1-H_2 en voz *creaky* ($r=,578$ y $p=,0001$). Esto significa que cuanto más desciende la frecuencia fundamental en la voz *creaky* frente a la modal, mayor es la diferencia negativa entre H_1-H_2 .

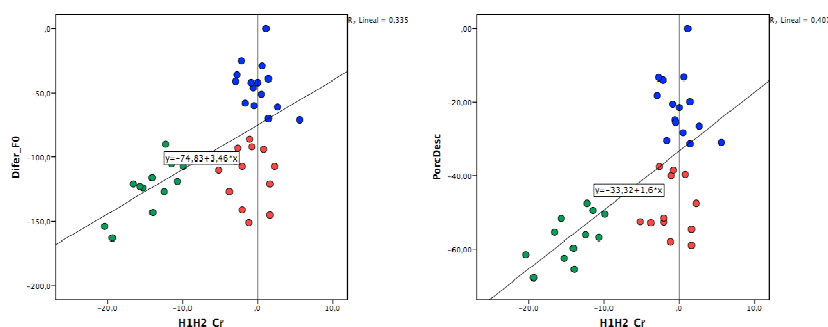


Figura 8. Gráficos de dispersión del descenso en la frecuencia fundamental relacionado con la diferencia de amplitud del primer y el segundo armónico (a la izquierda) y del porcentaje de descenso relacionado con la misma variable (a la derecha). Línea de regresión y línea de referencia al eje de abscisas (la línea representa la frontera entre los valores positivos y negativos de H_1-H_2).

En la figura 8 se observan tres grupos diferentes de ajustes laríngeos. Los grupos coloreados en verde y rojo suponen valores muy negativos en la variable de diferencia de F_0 creaky y modal; de hecho, los datos se sitúan por debajo de los 100 Hz. de descenso, es decir, el grado de tensión longitudinal es bajo. El grupo coloreado en verde refleja valores muy negativos en la diferencia de H_1-H_2 y los datos presentan una alta correlación visible en el gráfico de dispersión; esto quiere decir que su constricción laríngea es alta. Los grupos coloreados en rojo y en azul presentan dispersión de datos y diferencia de H_1-H_2 no tan fuertemente negativa e, incluso, positiva en algunos casos. En ambos casos, la constricción laríngea es más baja. Si se usa la variable de porcentaje de bajada con la diferencia de H_1-H_2 , la correlación resulta ser positiva, moderada y muy significativa ($r=,638$ y $p=,0001$). El gráfico (ver figura 8) es muy parecido al anterior. Los grupos rojo y verde se sitúan por debajo del porcentaje del 40% de descenso.

Los grupos son bastante homogéneos estadísticamente pero, al escuchar las muestras pertenecientes a cada grupo, se observa claramente que las hablantes recurren a diferentes ajustes laríngeos a la hora de producir las muestras.

5. DISCUSIÓN Y CONCLUSIONES

Las muestras seleccionadas para el presente estudio son solo aquellas que presentan F_0 . Así pues, las voces aperiódicas y las voces tensas, según la catalogación de Keating, Garellek y Kreiman (2015), se descartan por este motivo. El análisis acústico llevado a cabo muestra que, cuando se les pide producir una voz *creaky*, las hablantes utilizan como estrategia unánime el descenso en la frecuencia fundamental, es decir, la relajación de la tensión longitudinal de los pliegues vocales. En este caso, todas las hablantes siguen la estrategia de la voz *creaky* prototípica.

En cuanto a las diferencias entre H_1-H_2 y H_1-A_3 : 13 hablantes muestran valores negativos en ambos parámetros según el prototipo de voz *creaky*; otras 13 hablantes presentan diferencia negativa entre H_1-H_2 y positiva entre H_1-A_3 lo que implica un grado moderado de constricción laríngea lo que puede encajar con cualquiera de los tipos de voz *creaky* propuestos por Keating, Garellek y Kreiman (2015) a excepción del prototípico y del no constreñido; 5 muestras dan como resultado una diferencia positiva entre H_1-H_2 y negativa entre H_1-A_3 , lo que puede significar un modelo de voz *creaky* no constreñida si bien, el valor negativo de H_1-A_3 implica el refuerzo de la estructura armónica alta; por último, 7 muestras dan como resultado una diferencia positiva en ambos parámetros lo que implica una voz no constreñida por completo. En este último caso, sería necesario el análisis de otros parámetros tales como la irregularidad del pulso o los subarmónicos para corroborar acústicamente la existencia de voz *creaky*.

Los gráficos de dispersión han ofrecido también datos muy interesantes como la agrupación de las estrategias fonatorias basadas en el grado de tensión longitudinal de los pliegues vocales y el grado de constricción laríngea.

Aunque el presente estudio no analiza los parámetros de la voz modal, es destacable que una informante presente valores negativos para H_1-H_2 no esperables en este tipo de fonación y para lo que no se encuentra explicación. Un futuro análisis con algún programa diferente quizás contribuya a aclarar esta duda. Por otro lado, la diferencia H_1-A_3 es negativa en voz modal para 17 muestras (42,5%). Una posible explicación es que las voces modales de estas hablantes tengan un cierto grado de constricción o que las medidas acústicas registren un escape de aire, posiblemente debido a causas fisiológicas de la voz femenina. Se hace necesario un análisis perceptivo por parte de un experto en un futuro estudio.

Como se ha señalado en el apartado dedicado al estado de la cuestión, las estructuras epiglóticas tienen un papel generador de sonido que no puede ser

eludido, ya que participan activamente en los modos de fonación (Esling y Harris, 2005; Edmonson y Esling, 2006). Los resultados observados en ciertas informantes que no se han comportado de la manera considerada estándar y las referencias bibliográficas consultadas constituyen suficientes indicios como para abordar nuevas investigaciones en el futuro que incluyan estudios de otra índole que permitan profundizar en el conocimiento que se tiene sobre la voz *creaky*. Se hace necesario un análisis biomecánico que examine parámetros como la tensión y la masa del cuerpo y la cubierta de los pliegues vocales y un análisis acústico más profundo con herramientas especializadas en el análisis de las voces no modales como el VoiceSauce¹² que proporciona un número mayor de medidas como el jitter, el shimmer, el promedio de ruido interarmónico, el pico cepstral, etc. En este sentido, es muy interesante el análisis LTAS de un segmento de la muestra para observar no solo lo que ocurre en un corte espectral. Por último, se desprende del presente estudio la necesidad de realizar análisis por parte de expertos en patologías vocales y experimentos perceptivos para averiguar si los oyentes son capaces de distinguir los tipos de voz *creaky* que se han observado en el estudio.

AGRADECIMIENTOS: Gracias a Juana Gil, directora del Laboratorio de Fonética del CSIC, por permitir usar los datos de CIVIL para el presente estudio; a Helena Alves por facilitarme las grabaciones; a Patricia Infante por ayudarme con los primeros pasos del estudio; a José M^a Lahoz y Rafael Olalla por sus consejos estadísticos; a Pedro Gómez Vilda, catedrático de la Universidad Politécnica de Madrid, por su ayuda a la hora de analizar algunos segmentos conflictivos y a Nuria Polo por su apoyo incondicional, sus valiosos consejos y sus continuas correcciones.

6. REFERENCIAS BIBLIOGRÁFICAS

- ABERCROMBIE, D. (1967): *Elements of general phonetics*, Edimburgo, Edinburgh University Press.
- BECK, J. M. (2010): «Organic Variation of the Vocal Apparatus», en W. J. Hardcastle, J. Laver y F. E. Gibbon (eds.): *The Handbook of Phonetic Sciences*, Oxford, John Wiley & Sons, 2010, pp. 119-155.

¹² Se trata de una aplicación para Matlab que realiza un análisis acústico de la voz mucho más detallado que el programa Praat. Esta herramienta ha sido desarrollada por la Universidad de California en Los Ángeles (UCLA). La página web informativa es: <http://www.seas.ucla.edu/spapl/voicesauce/>

-
- BELOTEL-GRENIÉ, A. y M. GRENIÉ (2004): «The creaky voice phonation and the organization of Chinese discourse», en B. Bel (ed): *Proceedings of International Symposium on Tonal Aspects of Languages. With Emphasis on Tone Languages*, Pekín, The Institute of Linguistics in Chinese Academy of Social Sciences, pp. 5-8.
- BLOMGREN, M.; Y. CHEN, M. L, NG y H. R. GILBERT (1998): «Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers», *The Journal of the Acoustical Society of America*, 103, p. 2649.
- BOERSMA, P. y D. WEENINK (2013): Praat: doing phonetics by computer. Versión 5.3.56. <http://www.praat.org> [10/02/2015]
- CAMARGO, Z.; S. MADUREIRA, A. N. PESSOA y L. C. RUSILO (2012): «Voice quality and gender: some insights on correlations between perceptual and acoustic dimensions», en Q. Ma, H. Ding y D. Hirst (eds): *Abstract Book of 6th International Conference on Speech Prosody*, Shanghai, Tongji University Press, 1, pp. 115-118.
- CATFORD, J. C. (1964): «Phonation types: The classification of some laryngeal components of speech production» en D. Abercrombie, D. B. Fry, P. A. D. McCarthy, N. C. Scott y J. L. L. Trim (eds.): *In honour of Daniel Jones*, Londres, Longman, pp. 26-37.
- CATFORD, J. C. (1977): *Fundamental problems in phonetics*, Edimburgo, Edinburgh University Press, vol. 1, p. 977.
- CRYSTAL, D. (1976): *Prosodic systems and intonation in English*, Cambridge, Cambridge University Archives, vol. 1.
- COATES, J. (1986): *Women, Men, and Language: A Sociolinguistic Account of Sex Differences*, Londres, Longman.
- COBETA, I.; F. NÚÑEZ y S. FERNÁNDEZ (2013): *Patología de la voz*, Marge books, Madrid.
- CULLEN, A.; J. KANE, T. DRUGMAN y N. HARTE (2013): «Creaky voice and the classification of affect», en *Proceedings of WASSS (Workshop on Affective Social Speech Signals)*, Grenoble.
- DEJONCKERE, P. H.; P. BRADLEY, P. CLEMENTE, G. CORNUT, L. CREVIER-BUCHMAN, G. FRIEDRICH y V. WOISARD (2001): «A basic protocol for
-

- functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques», *European Archives of Oto-rhino-laryngology*, 258(2), pp. 77-82.
- DILLEY, L.; S. SHATTUCK-HUNAGEL Y M. OSTENDROF (1996): «Glottalization of word-initial vowels as a function of prosodic structure», *Journal of Phonetics*, 24, pp. 423-444.
- DRAKE, O. J. (1937): «Toward an improved vocal quality», *Quarterly Journal of Speech*, 23(4), pp. 620-626.
- EDMONSON, J. A. y J. H. ESLING (2006): «The valves of the throat and their functioning in tone, vocal register and stress: laryngoscopic case studies», *Phonology*, 23, 2, pp. 157-191.
- ERICKSON, D.; T. SHOCHI, C. MENEZES, H. KAWAHARA y K. I. SAKAKIBARA (2008): «Some non-F0 cues to emotional speech: an experiment with morphing», *Speech Prosody*, pp. 6-9.
- ESLING, J. H. (1983): «A laryngographic investigation of phonation type and laryngeal configurations», *Working Papers of the Linguistics Circle*, 3 (1), pp. 14-36.
- ESLING, J. y J. HARRIS (2005): «States of the glottis: An articulatory phonetic model based on laryngoscopic observations» en W. Hardcastle y J. Beck (eds.): *A Figure of Speech: A Festschrift for John Laver*, Nueva Jersey, Lawrence Erlbaum Associates, pp. 247-383.
- ESPOSITO, C. M. (2010): «Variation in contrastive phonation in Santa Ana del Valle Zapotec», *Journal of the International Phonetic Association*, 40, 2, pp. 181-198.
- FAWCUS, M. (1986): *Voice disorders and their management*, Chapman & Hall, Londres.
- FERNÁNDEZ-BAILLO GALLEGO DE LA SACRISTANA, R. (2013): *Índice acústico de discapacidad vocal (IADV) en población adulta: diseño de la escala, resultados y correlatos anatómico-fisiológicos*, tesis doctoral, Universidad Complutense, Madrid.
- FIELD, A. (2013): *Discovering statistics using IBM SPSS statistics*, Londres, Sage.

-
- FOUGHT, C. (2003): *Chicano English in context*, Nueva York, Palgrave Macmillan.
- FRIČ, M.; F. ŠRAM y J. G. ŠVEC (2006): «Voice registers, vocal folds vibration patterns and their presentation in videokymography», en E. Rajčan y E. Borsíková (eds): *Proceedings of ACOUSTICS High Tatras 06, 33rd International Acoustical Conference - EAA Symposium*, Eslovaquia, Štrbské Pleso, Slovak Acoustical Soc, pp. 42-45.
- FUJIMOTO, M. y K. Maekawa (2003): «Variation in phonation types due to paralinguistic information: An analysis of high-speed video images», en M., J. Solé, D. Recasens y J. Romero (eds): *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, Causal Productions, pp. 2401-2404.
- GERRATT, B. y J. KREIMAN (2001): «Toward a taxonomy of nonmodal phonation», *Journal of Phonetics*, 29, 4, pp. 365-381.
- GERRATT, B. y J. KREIMAN (2004): «Perceptual evaluation of voice quality», en R. D. Kent (ed.): *The MIT Encyclopedia of Communication Disorders*, Cambridge, Massachusetts, The MIT Press, pp. 78-80.
- GICK, B.; I. WILSON y D. DERRICK (2013): *Articulatory Phonetics*, Malden, Wiley-Blackwell.
- GIL FERNÁNDEZ, J. (1988): *Los sonidos del lenguaje*, Madrid, Síntesis.
- GIL FERNÁNDEZ, J. (2012): «La cualidad de voz y la comparación judicial de voces», presentación en *II Jornadas (In)formativas de Lingüística Forense*. Facultad de Filosofía y Letras, Universidad Autónoma de Madrid.
http://liceu.uab.es/~joaquim/phonetics/fon_prosod/suprasegmentales_fonacion.html [20/12/2015]
- GOBL, C. (1989): «A preliminary study of acoustic voice quality correlates», *STL-QPSR*, 4, pp. 9-21.
- GOBL, C. y A. NÍ CHASAIDE (1992): «Acoustic characteristics of voice quality», *Speech Communication*, 11 (4), pp. 481-490.
- GOBL, C. y A. NÍ CHASAIDE (2003): «The role of voice quality in communicating emotion, mood and attitude», *Speech communication*, 40, 1, pp. 189-212.
- GOBL, C. y A. NÍ CHASAIDE (2010): «Voice source variation and its communicative functions», en W. J. Hardcastle, J. Laver y F. E. Gibbon (eds.): *The Handbook of Phonetic Sciences*, Oxford, Wiley-Blackwell, 2010², pp. 378-423.
-

-
- GORDON, M. (2001): «Linguistic aspects of voice quality with special reference to Athabaskan», en S. Tuttle y G. Holton (eds): *Proceedings of the 2001 Athabaskan Languages Conference*, Los Ángeles, Ed. Alaska Native Language Center, pp. 163-178.
- GORDON, M. y P. LADEFOGED (2001): «Phonation types: a cross-linguistic overview», *Journal of Phonetics*, 29 (4), pp. 383-406.
- HAMMARBERG, B. (1999): «Voice research and clinical needs», *Folia phoniatria et logopaedica*, 52, 1-3, pp. 93-102.
- HAMMARBERG, B. y J. GAUFFIN (1995): «Perceptual and acoustic characteristics of quality differences in pathological voices as related to physiological aspects», en O. Fujimura y M. Hirano (eds.): *Vocal Fold Physiology, Voice Quality Control*, San Diego, Ed. Singular Pub. Group, pp. 283-303.
- HANSON, H. M. (1995): *Glottal characteristics of female speakers*, tesis doctoral, Harvard University, Cambridge, Massachussets.
- HANSON, H. M. y E. S. CHUANG (1999): «Glottal characteristics of male speakers: Acoustic correlates and comparison with female data», *The Journal of the Acoustical Society of America*, 106, 2, pp. 1064-1077.
- HANSON, H. M.; K. N. STEVEN, H. K. J. KUO, M. Y. CHEN y J. SLIFKA (2001): «Towards models of phonation», *Journal of Phonetics*, 29, 4, pp. 451-480.
- HEDELIN, P. y D. HUBER (1990): «Pitch period determination of aperiodic speech signals», en *Proceedings of Acoustics, Speech, and Signal Processing*, Albuquerque, Institute of Electrical and Electronic Engineers Service Center, pp. 361-364.
- HENRICH, D. N. (2006): «Mirroring the voice from Garcia to the present day: Some insights into singing voice registers», *Logopedics Phoniatrics Vocology*, 31, 1, pp. 3-14.
- HENTON, C. y A. BLADON (1988): «Creak as a sociophonetic marker» en L. Hyman y C. N. Li, (eds.): *Language, Speech, and Mind: studies in honour of Victoria Fromkin*, Londres, Routledge, pp. 3-29.
- HEWLETT, N. y J. M. BECK (2013): *An introduction to the science of phonetics*, Londres, Routledge.
-

-
- HIRANO, M. y D. M. BLESS (1993): *Videostroboscopic examination of the larynx*, San Diego, Singular Publishing Group.
- HIROSE, H. (1971): «Laryngeal Adjustments for Vowel Devoicing in Japanese: An Electromyographic Study», *Haskins Laboratories Status Report on Speech Research SR-28*, pp. 157-165.
- HOLLIEN, H. (1974): «On vocal registers», *Journal of Phonetics*, 2, pp. 125-143.
- HOLLIEN, H. y R. W. WENDAHL (1968): «Perceptual study of vocal fry», *The Journal of the Acoustical Society of America*, 43, p. 506.
- HOLMBERG, E. B.; R. E. HILLMAN, J. S. PERKELL, P. C. GUIOD y S. L. GOLDMAN (1995): «Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice», *Journal of Speech, Language and Hearing Research*, 38, 6, p. 1212.
- HONIKMAN, B. (1964): «Articulatory settings» en D. Abercrombie, D. B. Fry, P. A. D. MacCarthy, N. C. Scott y J. L. L. Trim (eds.): *In honor of Daniel Jones*, Londres, Longmans, pp. 73-84.
- ISELI, M.; Y. L. SHUE y A. ALWAN (2007): «Age, sex, and vowel dependencies of acoustic measures related to the voice source», *The Journal of the Acoustical Society of America*, 121, 4, pp. 2283-2295.
- JAVKIN, H. R.; N. ANTOÑANZAS-BARROSO e I. MADDIESON (1987): «Digital Inverse Filtering for Linguistic Research», *Journal of Speech, Language, and Hearing Research*, 30, 1, pp. 122-129.
- JIANG, D. N.; J. H. TAO y L. H. CAI (2002): «Voice Quality Analysis under the Pitch Effect», en *Proceedings of the International Symposium on Chinese Spoken Language Processing*, Taipei.
- JOHNSON, K. (1997): *Acoustic and Auditory Phonetics*, Malden, Wiley- Blackwell, 2012³.
- JOHNSTONE, T. y K. R. SCHERER (1999): «The effects of emotions on voice quality», en J. Ohala (ed): *Proceedings of the XIVth International Congress of Phonetic Sciences*, San Francisco, University of California, pp. 2029-2032.
- KEATING, P. A. y C. ESPOSITO (2006): «Linguistic voice quality», *UCLA Working Papers in Phonetics*, 105, pp. 85-91.
-

- KEATING, P.; J. KUANG, C. ESPOSITO, M. GARELLEK y S. KHAN (2012): «Multi-dimensional phonetic space for phonation contrasts», póster presentado en *LabPhon 13*, Stuttgart, Alemania.
- KEATING, P. A. y M. GARELLEK (2015): «Acoustic analysis of creaky voice», póster presentado en una sesión especial sobre voz *creaky* en el congreso anual de la *Linguistic Society of America*, Portland, Estados Unidos.
- KEATING, P. A.; M. GARELLEK y J. KREIMAN (2015): «Acoustic properties of different kinds of creaky voice», en M. Wolters, J. Livingstone, B. Beattie, R. Smith, M. MacMahon, J. Stuart-Smith y J. Scobbie (eds): *Proceedings of 18th International Congress of Phonetic Sciences*, Glasgow, International Phonetic Association.
- KIM, J. Y. (2013): *The Use of Creaky Voice by Spanish Heritage Speakers in the US*, póster presentado en el *Seventh Heritage Language Research Institute*, Chicago.
- KREIMAN, J.; D. VANLANCKER-SIDTIS y B. R. GERRAT (2008). «Perception of voice quality», en D. B. Pisoni y R. E. Remez (eds.): *Handbook of Speech Perception*, Oxford, Blackwell Publishing, pp. 338-362.
- KREIMAN, J. y D. SIDTIS (2011): *Foundations of voice studies: An interdisciplinary approach to voice production and perception*, Hoboken, Wiley-Blackwell.
- KREIMAN, J.; Y. L. SHUE, G. CHEN, M. ISELI, B. R. GERRAT, J. NEUBAUER y A. ALWAN (2012): «Variability in the relationships among voice quality, harmonic amplitudes, open quotient, and glottal area waveform shape in sustained phonation», *Journal of the Acoustical Society of America*, 132(4), pp. 2625-2632.
- LADEFOGED, P. (1971): *Preliminaries to linguistic phonetics*, Chicago, University of Chicago Press.
- LAVER, J. (1968): «Voice quality and indexical information», *International Journal of Language & Communication Disorders*, 3, 1, pp. 43-54.
- LAVER, J. (1975): «*Individual features in voice quality*», tesis doctoral, Universidad de Edimburgo.
- LAVER, J. (1980): *The Phonetic Description of Voice Quality*, Cambridge, Cambridge University Press.

-
- LAVER, J.; S. WIRZ, J. MACKENZIE y S. M. HILLER (1981): «A perceptual protocol for the analysis of vocal profiles», *Edinburgh University Department of Linguistics Work in Progress*, 14, pp. 139-155.
- LAVER, J. (1994): *Principles of phonetics*, Cambridge, Cambridge University Press.
- LUO, J. (2012): *Affective computing and intelligent interaction*, vol. 137, Springer Science & Business Media.
- MENDOZA-DENTON, N. (2011): «The semiotic hitchhiker's guide to creaky voice: circulation and gendered hardcore in a Chicana/o gang persona», *Journal of the Linguistic Anthropology*, 21, 2, pp. 261-280.
- MICHEL, J. F. (1968): «Fundamental frequency investigation of vocal fry and harshness», *Journal of Speech, Language, and Hearing Research*, 11 (3), pp. 590-594.
- MILLER, R. (2004): *Solutions for singers: Tools for performers and teachers*, Oxford, Oxford University Press.
- MOOSMÜLLER, S. (2007): «The influence of creaky voice on formant frequency changes», *International Journal of Speech Language and the Law*, 8, 1, pp. 100-112.
- MORRIS, R. y A. B. HARMON (2010): «Describing voice disorders», en J. S. Damico, N. Müller y M. J. Ball (eds.): *The handbook of language and speech disorders*, Malden, John Wiley & Sons, pp. 455-473.
- MOSER, H. M. (1942): «Symposium on Unique Cases of Speech Disorders; Presentation of a Case», *Journal of Speech and Hearing Disorders*, 7, 2, pp. 173-174.
- PERROT, P.; G. AVERSANO y G. CHOLLET (2007): «Voice disguise and automatic detection: review and perspectives», en Y. Stylianou, M. Faúndez-Zanuy y A. Esposito (eds): *Progress in nonlinear speech processing*, Springer Berlin Heidelberg, pp. 101-117.
- PODESVA, R. J. (2011): «Gender and the social meaning of non-modal phonation types», en C. Cathcart, I-H-Chen, G. Finley, S. Kang, C. S. Sandy y E. Stickles (eds): *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, Berkeley, Berkeley Linguistic Society, vol. 37, 1, pp. 427-448.
-

-
- REDI, L. y S. SHATTUCK-HUFNAGEL (2001): «Variation in the realization of glottalization in normal speakers», *Journal of Phonetics*, 29 (4), pp. 407-429.
- RODMAN, R. (1998): «Speaker recognition of Disguised Voices: A Program for research», en M. Demirekler, A. Saranli, H. Altincay y A. Paoloni (eds): *Proceedings of the Consortium on Speech Technology in Conjunction with the Conference on Speaker Recognition by Man and Machine: Directions for Forensic Applications*, Ankara, Publishing Arm, pp. 9-22.
- SAN SEGUNDO, E.; H. ALVES y M. FERNÁNDEZ TRINIDAD (2013): «CIVIL Corpus: voice quality for speaker forensic comparison», *Procedia-Social and Behavioral Sciences*, 95, pp. 587-593.
- SAPIR, E. (1927): «Speech as a personality trait», *American Journal of Sociology*, 32, 6, pp. 892-905.
- SCHERER, K. R. (1989): «Vocal correlates of emotional arousal and affective disturbance», en H. E. Wagner y A. E. Manstead (eds.): *Handbook of Social Psychophysiology*, Chichester, John Wiley & Sons, pp. 165-197.
- SCHERER, S.; J. KANE; C. GOBL y F. SCHWENKER (2013): «Investigating fuzzy-input fuzzy-output support vector machines for robust voice quality classification», *Computer Speech & Language*, 27(1), pp. 263-287.
- SEYFAHRT, S. y M. GARELLEK (2015): «Coda glottalization in American English», en M. Wolters, J. Livingstone, B. Beattie, R. Smith, M. MacMahon, J. Stuart-Smith y J. Scobbie (eds): *Proceedings of 18th International Congress of Phonetic Sciences*, Glasgow, International Phonetic Association.
- SHAW, F. y W. CROCKER (2015): «Creaky voice as a stylistic feature of young american female speech: an intraspeaker variation study of Scarlett Johansson», *Lifespans and Styles*, 1, pp. 21-27.
- SPSS Inc. Released (2008): *SPSS Statistics for Windows*, Version 17.0, Chicago, SPSS Inc.
- STEVENS, K. (1977): «Physics of laryngeal behavior and larynx modes», *Phonetica*, 34, pp. 264-279.
-

-
- STORY, B. H. y I. R. TITZE (2002): «A preliminary study of voice quality transformation based on modifications to the neutral vocal tract area function», *Journal of Phonetics*, 30 (3), pp. 485-509.
- SUNDBERG, J. (2013). «The perception of singing», en D. Deutsch (ed.): *The psychology of music*, Amsterdam, Academic Press, pp. 59-98.
- SVEČ, J. (2000): *On vibration properties of human vocal folds: voice registers, bifurcations, resonance characteristics, development and application of videokymography*, tesis doctoral, Universidad de Groningen.
- TAJFEL, H. (1974): «Social identity and intergroup behavior», *Social Science Information*, 13, pp. 65-93.
- TESHIGAWARA, M. (2003): «Voices in Japanese animation: a phonetic study of vocal stereotypes of heroes and villains in Japanese culture», tesis doctoral, Universidad de Victoria.
- TITZE, I. R. (1990): «Interpretation of the electroglottographic signal», *Journal of Voice*, 4 (1), pp. 1-9.
- TITZE, I. R. (1994). *Principles of Voice Production*, Englewood Cliffs, New Jersey. Prentice Hall.
- TRASK, L. R. (1996): *A dictionary of phonetics and phonology*, Londres, Routledge.
- YANUSHEVSKAYA, I.; C. GOBL y N.A. CHASAIDE (2005): «Voice quality and F₀ cues for affect expression: implications for synthesis», en *Proceedings of INTERSPEECH*, Lisboa, Curran Associates Inc, pp. 1849-1852.
- YUASA, I. P. (2010): «Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women?», *American Speech*, 85, pp. 315-337.