

Indicadores fonorrespiratorios de normalidad y patología en la clínica vocal

Josep Maria Vila-Rovira, Jesus Valero-Garcia, Laura Gonzalez-Sanvisens

Universitat Ramon Llull

Resumen

La disfonía es una alteración de las cualidades de la voz. Estas cualidades pueden ser medidas y los valores obtenidos pueden aportar información relevante para el profesional clínico de la voz. Esta investigación pretende identificar las medidas más válidas en el estudio objetivo de la voz patológica y establecer valores que identifiquen a la población disfónica. Se ha estudiado la voz de 141 personas con alteraciones vocales y se ha comparado con un grupo control de 99 sujetos. Los resultados obtenidos muestran diferencias significativas en diversos parámetros y se pueden perfilar valores de corte en la frecuencia máxima (537 Hz.- Do₄ en hombres y 667 Hz.- Mi₄ en mujeres) así como en el *Dysphonia Severity Index* (3.70). La diversidad y variabilidad de los datos observados reclaman al clínico la necesidad de integrar los datos obtenidos con las medidas objetivas con sus percepciones cualitativas, las valoraciones subjetivas del propio paciente y las informaciones obtenidas por las imágenes laríngeas.

Palabras clave: Calidad vocal; Disfonía; Exploración de la voz; Voz.

Abstract

Dysphonia is a disorder of the voice qualities. These qualities can be measured and these values may provide relevant information for clinical professionals of voice. This research aims to identify the most valid measures in the study of pathological voice and set cut-off values to identify the population hoarse. We have studied the voice of 141 people with vocal disorders and compared with a control group of 99 subjects. The results show significant differences in several parameters and values can shape cut-off in the maximum frequency (537 Hz - C₄ in men and 667 Hz - Mi₄ in women) and in the Dysphonia Severity Index (3.70). The diversity and variability observed of the data call to the clinician to integrate the data the objective measures with their qualitative perceptions, subjective assessments of the patient and the information obtained by the laryngeal images.

Keywords: Dysphonia; Voice; Voice assessment; Voice quality.

Introducción

La disfonía es un fenómeno multidimensional, que afecta a las cualidades de la voz del hablante, que es percibida por sus interlocutores y que altera la vida cotidiana de quien la padece. Aronson (1980) la define como la anomalía de la voz percibida por aquel que la escucha y que afecta sea a la frecuencia, a la intensidad, al timbre, a la ligereza o bien

a una combinación entre ellas. El clínico y el investigador precisan de instrumentos que permitan evidenciar los cambios que la voz sufre y éstos deben responder a las diversas dimensiones que la voz y su uso comunicativo presentan. En este aspecto, coincidimos con las diversas propuestas de abordaje multi-instrumental de la disfonía. Partiendo del amplio acuerdo obtenido por el protocolo elaborado por el Comité de Foniatría de la Sociedad Europea de Laringólogos (Dejonckere *et al.* 2001), recogiendo los últimos trabajos de revisión (Carding, Wilson, MacKenzie y Deary, 2009; Oates, 2009) y sin olvidar las propuestas metodológicas e instrumentales que han aparecido en nuestro contexto territorial (Arias, 1992; Rodríguez-Parra, Adrián y Casado, 2009), centraremos nuestra atención en algunas medidas de la función vocal, algunas muy habituales en la práctica clínica y otras más novedosas. En todos estos referentes se determina la necesidad de establecer unos acuerdos de mínimos para el diagnóstico de las alteraciones de la voz y disponer de herramientas comparables para ser usadas en las investigaciones sobre la eficacia de los tratamientos y de las nuevas técnicas de evaluación. En estos protocolos y revisiones se tienen en cuenta las informaciones aportadas por las pruebas de imagen laríngea, por las valoraciones perceptivas, por los datos del examen acústico, por las medidas aerodinámicas y por las valoraciones del hándicap vocal.

En primer lugar, el proceso diagnóstico precisa de la observación directa de la laringe. En este aspecto, Dejonckere *et al.* (2001) afirman que la videolaringoestroboscopia es el principal instrumento para el diagnóstico etiológico de las alteraciones de la voz. Se trata de la visualización y grabación de los pliegues vocales en acción fonatoria con el apoyo de una fuente de luz sincronizada con la vibración glótica. Esta prueba permite valorar también la calidad de la vibración de los pliegues vocales y puede ser utilizada para la evaluación de la efectividad de los tratamientos médicos o quirúrgicos. Se afirma que esta pertenencia está basada en la buena reproducibilidad intra e interobservador, en la no redundancia de la información aportada y en la relación que se establece entre lo observado y la fisiología vocal. Sin embargo, en nuestro país la videolaringoestroboscopia no es, todavía, la prueba diagnóstica para todos los pacientes afectados por alteraciones de la voz. Compartimos la importancia de la captación de las imágenes laríngeas tanto para el proceso diagnóstico como para la mejor comprensión de la producción vocal tanto en laringes sanas como en las de las personas con disfonía. En nuestra investigación hemos centrado nuestra atención en otras perspectivas para la medida y valoración del

fenómeno vocal, motivo por el cual no profundizamos en la descripción y análisis de los procedimientos de captación de imágenes laríngeas.

En segundo lugar, debemos hacer referencia a la desarrollo de instrumentos para la valoración de las consecuencias que la disfonía provoca en quien la padece. A pesar de que nuestro trabajo no se dedica a ellas, consideramos necesario no perder de vista la mirada subjetiva del paciente. Entendemos que ceder la palabra al paciente para opinar sobre su enfermedad enriquece y complementa la visión objetiva y subjetiva del clínico. Siguiendo a Zraick y Risner (2008), podemos identificar la existencia de hasta doce cuestionarios de autopercepción de la disfonía. La descripción y análisis de todas estas herramientas supera ampliamente el objetivo de este documento y por este motivo preferimos referirnos únicamente al *Voice Handicap Index* (Jacobson *et al.*, 1997), por ser el único traducido, adaptado y validado a la lengua española (Núñez-Batalla *et al.* 2007). Se trata de un cuestionario compuesto por treinta ítems, en el que el paciente expresa con qué frecuencia sufre algunas limitaciones en la vida cotidiana motivadas por su voz. Siete años después de su publicación, se revisó y se propuso una versión corta de sólo diez ítems y se argumentó que no hay ningún beneficio en el uso de la versión completa (Rosen, Lee, Osborne, Zullo y Murry, 2004). La aceptación de esta herramienta ha sido muy importante. Existen diversos estudios que avalan su fiabilidad y validez y ha sido empleada en numerosos trabajos científicos como contraste de otros instrumentos de medida de la disfonía (Deary, Webb, Mackenzie, Wilson y Carding, 2004; Steen *et al.* 2008; Webb *et al.* 2007). Un reciente estudio de revisión presenta el VHI-10 como uno de los instrumentos más robustos para la valoración de la autopercepción de la disfonía (Carding *et al.* 2009). El *Voice Handicap Index* ha sido traducido a muchas lenguas (Verdonck-de Leeuw *et al.* 2008) y en nuestro país hace algunos años que se emplean las dos versiones del VHI, pero no ha sido hasta 2007 que se ha publicado una versión castellana adaptada y validada.

El tercero de los ámbitos de la exploración para el diagnóstico de las alteraciones de la voz es la valoración perceptiva, es decir, la valoración que el clínico realiza de la voz del paciente con su sentido de la audición. Todos los protocolos de evaluación actuales incluyen esta dimensión y los estudios de revisión más recientes no olvidan esta herramienta de valoración clínica que es la escucha del paciente (Behrman, 2005; Dejonckere *et al.* 2001; Oates, 2009; Rodríguez-Parra *et al.* 2009; Webb *et al.* 2004). Para valorar perceptivamente la calidad vocal, disponemos de un referente innegable en la escala GRBAS (Hirano, 1981). Hirano definió una escala de valoración perceptiva

que responde al acrónimo GRBAS, donde la G (*Grade*) corresponde al grado de disfonía, la R (*Roughness*) se refiere a la raspositat o aspereza, la inicial B (*Breathiness*) valora la presencia ruido de aire en la voz, voz soplada, la debilidad o astenicitat tímbrica se valora con la inicial A (*Asthenics*) y, finalmente, se puntúa la característica de la constricción en el parámetro S (*Strained*). El clínico puntúa el nivel de severidad de la disfonía en una escala de cuatro puntos. Donde al primero se le otorga el valor 0 que responde a la ausencia de disfonía y los valores 1, 2 y 3 indican los niveles ligero, moderado y severo en cada uno de los parámetros valorados. Se trata de una escala perceptiva que es conocida y empleada de forma habitual en todo el mundo a pesar de que en nuestro país no todos los profesionales ORL y logopedas la usen aún. En nuestro estudio consideramos, siguiendo a Oates (2009), que se trata de la prueba estándar para determinar la validez de otros instrumentos de medida de la calidad vocal.

Siguiendo con las dimensiones de valoración de la disfonía, debemos referirnos seguidamente a las medidas aerodinámicas. La producción de la voz precisa del aire pulmonar para su generación. Por este motivo, la respiración ha tenido un papel destacado tanto en las valoraciones de la voz como en los métodos de mejora o tratamiento de la voz. Siguiendo las orientaciones de la Sociedad Europea de Laringólogos (Dejonckere *et al.* 2001) el indicador aerodinámico de la voz más simple es el Tiempo Máximo de Fonación (TMF). Se trata de la prolongación de la emisión de la vocal /a/ durante el máximo tiempo posible en una frecuencia e intensidad confortable tras una inspiración profunda (Hirano, 1981; Neiman y Edeson, 1981). Se recomienda realizar una demostración y tres ensayos y se considera que puede tener una cierta sensibilidad al entrenamiento y a la fatiga. Son diversos los estudios que presentan datos normativos sobre el tiempo máximo de fonación y los valores que aportan presentan una alta variabilidad (Hakkesteeft, Brocaar, Wieringa y Feenstra, 2006; Rodríguez-Parra *et al.* 2009; Wuyts *et al.* 2000). Finalmente, nos parece idóneo referirnos al reciente trabajo presentado por Speyer *et al.* (2010) en el que ponen de manifiesto que la medida del tiempo máximo de fonación es una medida fiable y que no se ve afectada por el número de ensayos que se realicen ni de si se realiza en diferentes días o por diversos observadores.

Prosiguiendo con la descripción y análisis de las medidas aerodinámicas, nos queremos referir ahora al valor del Tiempo Máximo de Espiración (TME). No se trata de un dato directamente relacionado con la fonación pero nos permitirá acceder a un parámetro de relación entre fonación y espiración. El TME se obtiene pidiendo al

paciente que prolongue un soplo espiratorio durante el máximo tiempo posible con la emisión de un fonema /s/. En el mismo sentido recientemente, un grupo de Andalucía formado por Rodríguez-Parra, *et al.* (2009) en su protocolo multidimensional para la evaluación de la disfonía llamado Teatinos, proponen la anotación del TME y informan de diferencias significativas entre los valores obtenidos entre sujetos con afectación vocal y sin ella, así como entre los valores obtenidos antes y después de una intervención terapéutica. La relación entre el TME y el TMF se recoge en el denominado cociente s/z. Se interpreta como un indicador de la eficiencia de cierre glótico. Como nos presentan Gelfer y Pazera (2006), hay diversidad de datos no concluyentes sobre la fiabilidad de esta medida y sugieren que el uso clínico de la misma debe realizarse con cautela.

Finalmente, en el análisis de las dimensiones que forman parte del protocolo debemos referirnos a las medidas objetivas de calidad vocal. En este sentido, cabe decir que los avances de la informática han puesto al alcance de clínicos e investigadores distintas aplicaciones que aportan información sobre el sonido de la voz. El análisis acústico de la voz es una mirada no invasiva al fenómeno vocal que permite objetivar la evaluación de la voz con parámetros numéricos. En la tradición clínica los valores habitualmente observados son la frecuencia fundamental espontánea, las frecuencias mínima y máxima y las intensidades mínima y máxima. Por otra parte, gracias a los análisis informatizados de la voz es posible obtener diversos valores que expresan las características de una onda sonora. Mediante la transformada de Fourier es posible calcular múltiples valores. La explicación de todos y cada uno de los parámetros que se pueden calcular con los diferentes programas rebasa ampliamente los objetivos de este trabajo. Hay que referirse en este ámbito al trabajo de Baken (1999) donde se recogen las bases, datos y utilidades de estos parámetros acústicos.

Tomamos de nuevo la referencia del protocolo básico de la Sociedad Europea de Laringólogos (Dejonckere *et al.* 2001) donde se considera que las medidas de perturbación de la amplitud y de la frecuencia (*shimmer* y *jitter*) y las de relación entre la señal y el ruido son las más robustas para cuantificar las características perceptuales de la calidad vocal. Se señala, no obstante, que una de las limitaciones que presenta en ese momento el análisis acústico informatizado es la incapacidad de analizar señales acústicas fuertemente aperiódicas. Concretamente, recomienda no considerar los valores de perturbación de amplitud y periodo en porcentajes que superen el 5%. En referencia a los índices de señal / ruido considera que no hay suficientes estudios de

estandarización. Además, el protocolo básico de 2001 propone el uso de algunos valores del fonetograma (representación conjunta de intensidad y frecuencia que delimita el campo vocal de un hablante). Considera que realizar el fonetograma completo es muy costoso en cuanto al tiempo dedicado y propone que se tomen las tres medidas más sensibles al cambio: intensidad mínima y frecuencias máxima y mínima. Relacionado con los valores de período e intensidad, nos parece relevante uno de los últimos estudios publicados (Ma *et al.* 2007) donde se comparan los valores del fonetograma de un grupo de noventa mujeres con patología laríngea con otro de treinta y cinco mujeres sin patología vocal. Podemos destacar que las diferencias entre los dos grupos se manifiestan significativas para todos los parámetros estudiados, frecuencia mínima y máxima, intensidad mínima y máxima, rangos de frecuencia e intensidad y área vocal (producto del rango de frecuencias y de intensidades en semitonos).

La diversidad de medidas acústicas que pueden ofrecer los programas informáticos de análisis vocal así como la discutida fiabilidad y validez de las mismas ha llevado a diversos grupos de investigadores a la búsqueda de índices que agrupen las medidas más robustas y que puedan sintetizar el concepto de calidad vocal o de severidad de la disfonía. Podemos señalar la existencia y desarrollo de varios de éstos índices. Para nuestro estudio nos ha interesado centrarnos en el *Dysphonia Severity Index* (DSI) de Wuyts *et al.* (2000) por ser el que más desarrollo científico tiene hasta el momento. El DSI combina de manera ponderada los datos objetivos más significativos de una exploración vocal acústica y aerodinámica. El DSI recoge el tiempo máximo de fonación, la frecuencia fundamental más aguda, la intensidad más baja y el *jitter* en porcentaje. Su ponderación se expresa en la fórmula siguiente: $DSI = 0.13 \times TMF + 0.0053 \times F_0 \text{ máx.} - 0.26 \times \text{Intensidad mín.} - 1.18 \times \text{jitter} + 12.4$. En los últimos años, ha aumentado el uso de este indicador y ha sido estudiada su variabilidad en función de la edad y el género, interobservador, en test-retest y como indicador de cambios tras el tratamiento logopédico y la fonocirugía. (Hakkesteegt *et al.* 2006; Hakkesteegt, Wieringa, Brocaar, Mulder y Feenstra, 2008; Hakkesteegt, Brocaar y Wieringa, 2010). La experiencia de Wuyts *et al.* (2001) con el DSI ha sido seguida por varios grupos de investigadores en la búsqueda de nuevos índices en los que se combinaran varios parámetros acústicos que hicieran posible una valoración objetiva con la mejor capacidad de categorización, máxima sensibilidad al cambio, reproducibilidad y correlación con las percepciones de los jurados de escucha. En esta línea hay que tomar en consideración los trabajos liderados por Awan y por Maryn y sus colaboradores

(Awan y Roy, 2006; Awan y Roy, 2009; Awan, Roy y Dromey, 2009; Awan, Roy, Jette, Meltzner y Hillman, 2010; Maryn, De Bodt y Roy, 2010) que aportan principalmente el uso de muestras de habla continua combinada con las vocales sostenidas y la incorporación de medidas relacionadas con el espectro y el cepstrum, evitando el uso del *jitter* al que consideran que no es un indicador fiable extraíble de las vocales sostenidas.

En consecuencia de los datos hasta aquí aportados, los objetivos que nos proponemos en la investigación que presentamos tratan sobre la determinación de la significación de varias medidas aerodinámicas y acústicas en relación a la patología vocal y al posible establecimiento de puntuaciones de corte que sirvan al clínico de orientación en el proceso multiprofesional de la evaluación y diagnóstico de las afecciones vocales.

Método

Participantes

Para este estudio, la muestra del grupo experimental se seleccionó entre los pacientes diagnosticados por médicos ORL y atendidos, por alteraciones de la voz, en varios centros logopédicos con concierto económico con la administración sanitaria pública. El sistema de derivación de pacientes a los centros concertados no asegura, con fiabilidad, la categorización de los diagnósticos lesionales. Las muestras del grupo control se recogieron entre alumnado universitario, cantores de coros no profesionales, maestros participantes en cursos de educación de la voz y otros voluntarios convocados a tal efecto. Se excluyeron del grupo experimental los pacientes afectados de alteraciones de origen neoplásico, neurológico o de presbifonía. En el grupo control se excluyeron todas las personas que manifestaban haber sufrido problemas vocales con o sin diagnóstico confirmatorio. Después de una selección aleatoria, el grupo experimental se formó con un total de 141 pacientes y el grupo control con 99 personas. El conjunto de la muestra estaba formada por 38 hombres y 202 mujeres y la media de edad del grupo experimental era de 44 años y la del grupo control de 30. La diferencia de las medias de edad entre los dos grupos se mostró significativa ($z = 6.890$; $p < .01$).

La participación en este estudio, tanto en el grupo experimental como en el grupo control, ha sido voluntaria y con la firma de un documento de consentimiento informado. Las pruebas administradas son habituales en la práctica clínica de nuestro entorno y no se conocen estudios que informen de la nocividad de ninguna de ellas. Se

ha garantizado el anonimato de todos los datos registrados con una referencia alfanumérica específica para este estudio.

Materiales y procedimiento

Entre los meses de abril y diciembre de 2010, se procedió a la recogida de muestras mediante sesiones individuales de treinta minutos realizadas en espacios no insonorizados pero con un nivel sonoro inferior a 50 dB (Dejonckere *et al.* 2001). Las tareas solicitadas a cada participante pueden encontrarse en la tabla 1.

Tabla 1. Tareas para la recogida de muestras de voz

Decir el número de referencia de la investigación y la fecha del día
Producir los sonidos /s/, /z/ y /a/ durante el máximo de tiempo posible (dos intentos)
Producir la vocal /a/ en la tonalidad e intensidad que sea confortable (tres intentos de tres segundos como mínimo)
Recitar los días de la semana y alargar la vocal tónica /á/ del sábado (un intento de tres segundos como mínimo)
Producir un sonido en forma de sirena ascendente hasta el tono más agudo posible (tres intentos)
Producir tres sonidos a la mínima intensidad posible (tres intentos de tres segundos como mínimo)
Lectura de un texto

En relación a la muestra de voz en lectura, queremos indicar que no existe en nuestro ámbito territorial ningún texto que se utilice de manera consensuada para la exploración de la voz ni en investigaciones relacionadas con ella. Se dispuso la redacción de un escrito para esta finalidad que respondiera a ciertos criterios. La lectura usada es una adaptación de un cuento de Pere Calders, titulado “La mosca”. Dicho texto, tanto en su versión castellana como catalana contiene todos los sonidos consonánticos, vocálicos y diptongos de nuestras lenguas, palabras que comienzan en todas las vocales posibles y las curvas de entonación más usuales (interrogación, admiración, condición, subordinación...). La lectura está formada por fragmentos narrativos, descriptivos y diálogos y contiene alguna frase muy larga para valorar la gestión del aire en fonación. El texto puede ser leído en pocos minutos y no contiene palabras poco comunes o de pronunciación difícil. Ha sido probado por un grupo de

lectores voluntarios para comprobar su facilidad de lectura y la posibilidad de obtener datos como la frecuencia fundamental, el rango de frecuencias, los puntos de toma de aire, las curvas de entonación, los ataques vocales, entre otros. En nuestra investigación se ha utilizado para la valoración perceptiva realizada por un jurado de escucha experto compuesto por seis logopedas especializados en alteraciones de la voz que no presentan problemas auditivos conocidos. Los jueces han valorado la totalidad de las muestras vocales siguiendo los procedimientos habituales de la escala GRBAS de Hirano (1981) y se ha utilizado únicamente el parámetro G, por ser el más global de los cinco.

Las muestras vocales han sido recogidas por un micrófono de diadema AKG 555 (de respuesta plana e hipercardioides), un grabador digital Samson Zoom H4. Para la medida de la intensidad vocal se ha utilizado un sonómetro digital PCE 322-A. Las grabaciones se han realizado en formato digital *wave* con una frecuencia de muestreo de 44.100 Hz. a 16 bits y en mono. Las muestras de voz han sido analizadas en el Laboratorio de Logopedia de FPCEE Blanquerna de la Universitat Ramon Llull con los programas *Adobe Audition 2.0* de *Adobe Systems Incorporated* y *Multi Dimensional Voice Program* de *Kay Elemetrics*. Con estos programas se han obtenido los siguientes datos de cada muestra: Tiempo Máximo de Espiración (TME); Tiempo Máximo de Fonación con la consonante /z/ (TMF z); Tiempo Máximo de Fonación con la vocal /a/ (TMF a); Cociente TMFs/TMFz (Q s/z); Frecuencia fundamental promedio en habla espontánea a partir de la presentación y fecha (X F₀ espo.); Frecuencia fundamental en la emisión del fonema /z/ (F₀ /z/); Frecuencia fundamental en la emisión del fonema /a/ (F₀ /a/); Frecuencia fundamental en la emisión del fonema /á/ de /sábado/ (F₀ /á/ larga); Frecuencia fundamental máxima F₀ máx.); Intensidad mínima (Int. mín.) y *jitter* promedio de la /a/ sostenida (*jitter*). A partir de algunos de estos datos se ha calculado el DSI (Wuyts *et al.* 2001).

Los datos obtenidos se han tratado con el paquete estadístico PASW, versión 18.0, con él se han calculado estadísticos descriptivos y de significación así como se ha procedido al cálculo de curvas ROC (*Receiver Operating Characteristics*) para el establecimiento de una puntuación de corte. Este procedimiento estadístico permite valorar la fiabilidad de una prueba mediante el parámetro del área bajo curva (AUC) y establecer un punto de corte en función de la sensibilidad y la especificidad deseadas (Burgueño, García-Bastos y González-Buitrago, 1995).

Resultados

Presentaremos los resultados obtenidos organizados por bloques, iniciando por los datos relativos a los tiempos de espiración y fonación y al cociente s/z. Los resultados promedios se encuentran en la tabla 2.

Tabla 2. Valores promedio de Tiempo Máximo de Espiración (TME); Tiempo Máximo de Fonación con la consonante /z/ (TMF z); Tiempo Máximo de Fonación con la vocal /a/ (TMF a); Cociente TMFs/TMFz (Q s/z)

	TME	TMF z	TMF a	Q s/z
Total	13	12	11	1.14
Control	14	14	13	1.06
Experimental	12	11	10	1.21
Hombre	15	15	14	1.08
Mujer	13	12	11	1.16

Datos en expresados en segundos.

En el conjunto de los datos no se han asumido las condiciones de normalidad e igualdad de varianzas y se ha procedido a contrastar los resultados mediante pruebas no paramétricas. Se han encontrado diferencias estadísticamente significativas en los valores TMF en /z/ y en /a/ ($z = -4.272$ y -6.366 ; $p < .01$), así como en el Q s/z ($z = -2.496$; $p < .05$) entre los grupos control y experimental, siendo también significativa la diferencia de los tiempos de fonación en /z/ y en /a/ entre hombres y mujeres. ($z = -3.299$ y -3.323 ; $p < .01$). Se ha calculado en estadístico área bajo curva ROC (AUC) del cociente s/z y ha presentado un valor .594 que es considerado bajo, lo que nos apunta que no puede ser considerado como una prueba de valor discriminante en el proceso diagnóstico.

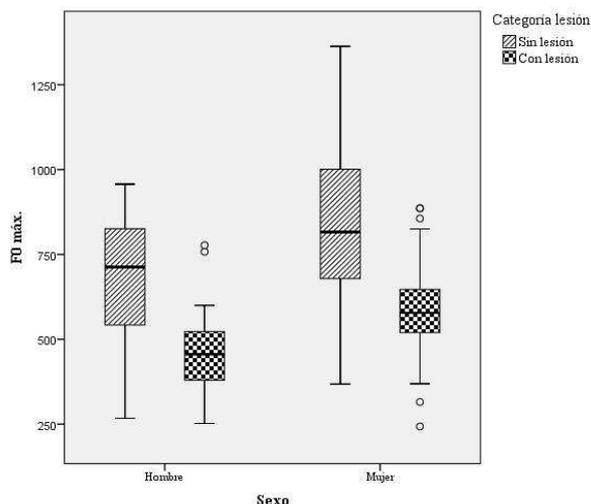
En referencia a los resultados a las frecuencias fundamentales de la voz, debemos diferenciar los resultados de hombres y mujeres. Los resultados obtenidos en este ámbito se encuentran resumidos en la tabla 3. Así mismo, podemos observar en el gráfico 1 que sólo se presentan diferencias significativamente diferentes en los valores relativos a la F_0 máxima (hombres: $z = -3.132$; $p < .01$ y mujeres: $z = -8.266$; $p < .01$).

Tabla 3. Valores en Hz. de las distintas frecuencias fundamentales

		X F ₀ esp.	F ₀ /z/	F ₀ /a/	F ₀ /á/ larga	F ₀ máx.
Hombres	Control	125	119	117	119	671
	Experimental	128	119	125	125	459
Mujeres	Control	206	175	189	199	827
	Experimental	196	178	194	200	580

Frecuencia fundamental promedio en habla espontánea a partir de la presentación y fecha (X F₀ espo); Frecuencia fundamental en la emisión del fonema /z/ (F₀ /z/); Frecuencia fundamental en la emisión del fonema /a/ (F₀ /a/); Frecuencia fundamental en la emisión del fonema /á/de /sábado/ (F₀ /á/ larga); Frecuencia fundamental máxima F₀ máx); Intensidad mínima (Int mín); *Jitter* promedio de la /a/ sostenida (*Jitter*).

Gráfico I: Valores medios de la frecuencia fundamental máxima para hombres y mujeres del grupo control y experimental en Hz.



Este hecho se hace más evidente cuando observamos la misma tabla expresando los datos en notas musicales y por lo tanto, más próximos a la escala de percepción humana de los tonos (Ver tabla 4).

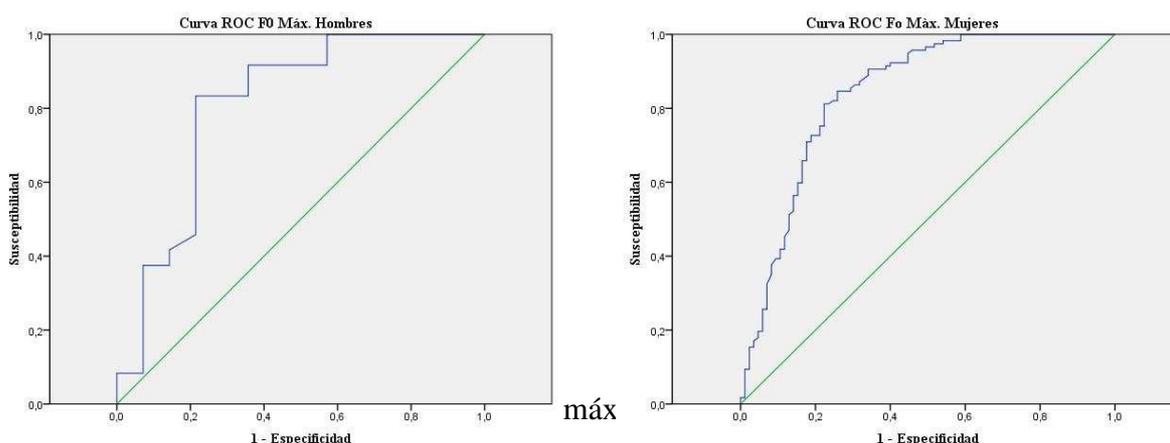
Tabla 4. Valores en notación musical de las distintas frecuencias fundamentales

		F ₀ X espo.	F ₀ /z/	F ₀ /a/	F ₀ /á/ larga	F ₀ máx.
Hombres	Control	Si ₁	Sib ₁	Sib ₁	Sib ₁	Mi ₄
	Experimental	Do ₂	Sib ₁	Si ₁	Si ₁	Sib ₃
Mujeres	Control	Sol# ₂	Fa ₂	Fa# ₂	Sol ₂	Lab ₄
	Experimental	Sol ₂	Fa ₂	Sol ₂	Sol ₂	Re ₄

Frecuencia fundamental promedio en habla espontánea a partir de la presentación y fecha (X F₀ espo.); Frecuencia fundamental en la emisión del fonema /z/ (F₀ /z/); Frecuencia fundamental en la emisión del fonema /a/ (F₀ /a/); Frecuencia fundamental en la emisión del fonema /á/de /sábado/ (F₀ /á/ larga); Frecuencia fundamental máxima F₀ máx.).

La significación observada en las diferencias entre los valores promedio de la frecuencia máxima de los grupos control y experimental nos ha impulsado a calcular la curva ROC (ver gráfico 2) y obtener una nota de corte de 537 Hz. (Do₄), con una sensibilidad de .83 y una especificidad de .78 para los hombres con un AUC = .808. En el caso de las mujeres, la AUC = .841 y una nota de corte de 667 Hz. (Mi₄), con una sensibilidad de .81 y una especificidad de .78.

Gráfico 2. Curvas ROC para F₀ máx en hombres y mujeres



En el ámbito de las medidas de altura vocal, nos ha interesado observar si los valores de las frecuencias fundamentales obtenidos en las diversas tareas son equivalentes. Concretamente hemos comparado, con la prueba de los signos de

Wilcoxon, si las F_0 obtenidas con la media de frecuencias en la tarea de presentación de cada participante respecto a las obtenidas en la emisión de la consonante /z/, en la emisión de las vocales /a/ sostenidas y, finalmente, en la prolongación de la /á/ de sábado en el recitado de los días de la semana eran equivalentes. Los resultados nos muestran que existen diferencias entre la F_0 promedio espontánea y la F_0 /z/ y F_0 /a/. En cambio aparecen como equivalentes los valores obtenidos entre la F_0 promedio espontánea y la F_0 de la /a/ prolongada de /sábado/ (ver tabla 5).

Tabla 5. Prueba de los rangos con signo de Wilcoxon entre los distintos procedimientos de obtener la F_0

	F_0 /z/ - F_0 X espo.	F_0 /a/ - F_0 X espo.	F_0 /á/ larga - F_0 X espo.
Z	-11.060	-5.947	-1.143
Sig.asintót. (bilateral)	.000	.000	.253

El tercer grupo de resultados analizados corresponde a los dos últimos parámetros que forman parte del *Dysphonia Severity Index* y al propio DSI. En relación con las medidas de la intensidad mínima (Int. Mín.), el *jitter* promedio de la /a/ sostenida (Jitter) y el DSI nuestros datos arrojan diferencias significativas entre el grupo control y el grupo de pacientes con alteración vocal documentada ($z = -4.392$, $p < .01$; $z = -2.362$, $p < .05$; $z = -9.320$, $p < .01$, respectivamente). Los valores medios pueden encontrarse en la tabla 6.

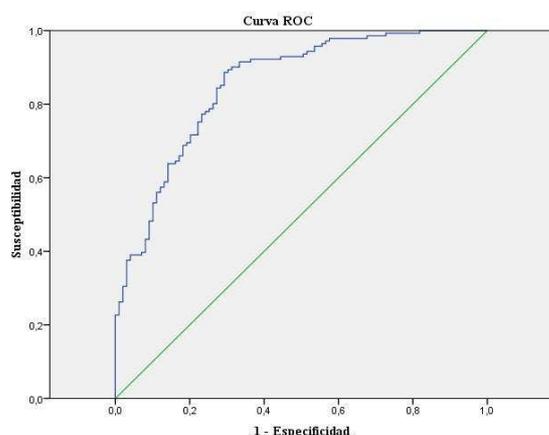
Tabla 6. Valores medios para la intensidad mínima en dB, el *jitter* en porcentaje y el *Dysphonia Severity Index*

	Int. Mín.	Jitter	DSI
Control	47 (4.2)	1.110 (.71)	5.1 (1.8)
Experimental	50 (4.1)	1.380 (.92)	2.2 (2.0)

Entre paréntesis, las desviaciones típicas.

En el cálculo de las curvas ROC, los resultados han mostrado para las variables intensidad mínima y *jitter* unas áreas bajo curva bajas (.666 y .590, respectivamente) que se consideran bajas y que por tanto no permiten la determinación de puntuaciones de corte que discriminen con precisión probabilística mínima los sujetos con alteraciones de los que no las presentan. Sin embargo, el índice DSI, calculado a partir de la combinación del tiempo de fonación, la frecuencia máxima, la intensidad mínima y el *jitter*, ha presentado un valor AUC de .853, considerado como útil (Burgueño *et al.*, 1995) y situando una nota de corte en los 3.70 puntos, con una sensibilidad de .77 y una especificidad de .77 (ver gráfico 3).

Gráfico 3. Curva ROC para DSI



Finalmente se ha calculado la correlación entre las distintas variables significativas y las puntuaciones que los jueces de escucha impusieron a las muestras vocales de todos los sujetos participantes. Entre los cinco parámetros de la escala GRBAS de Hirano (1999) se ha utilizado únicamente el parámetro global G que ha presentado un coeficiente de correlación intraclass (ICC) de .909 ($p < .001$), considerado excelente. Los datos obtenidos en nuestra investigación presentan correlaciones significativas ($p < .01$) entre el parámetro G y las variables relativas a los tiempos de fonación, al cociente *s/z*, así como con las variables frecuencia máxima, intensidad mínima, *jitter* y DSI (ver tabla 7).

Tabla 7. Correlaciones de las distintas variables con la puntuación media del parámetro G del GRBAS

		TME	TMF z	TMF a	Q s/z
Rho de Spearman					
Media G	Coef. de correlación	-.175	-.364	-.465	.194
	Sig. (bilateral)	.007	.000	.000	.003
		F ₀ X espo.	F ₀ /z/	F ₀ /a/	F ₀ /á/ larga /
Media G	Coef. de correlación	-.070	.110	.124	.104
	Sig. (bilateral)	.278	.090	.055	.107
		F ₀ Máx.	Int. Mín.	Jitter	DSI
Media G	Coef. de correlación	-.535	.388	.355	-.727
	Sig. (bilateral)	.000	.000	.000	.000

Discusión y conclusiones

El objetivo de este estudio era determinar la existencia de significación en las diferencias entre algunos valores respiratorios y fonatorios usados habitualmente en la exploración de las alteraciones de la voz. En aquellas variables que presentaran valores distintos entre el grupo de personas con patología vocal y los sujetos sanos, se proponía la determinación de puntos de corte que orienten al clínico sobre la existencia de estas alteraciones vocales.

Entre los resultados mostrados hasta ahora debemos analizar en primer lugar los relativos a los tiempos de espiración y fonación. Dichos resultados se muestran una tendencia a presentar valores más bajos que los publicados en la literatura anterior. Así, Rodríguez-Parra *et al.* (2009) informa de una media en el TME de 30 segundos para el grupo control y de 17 para el experimental, mientras que nuestros datos se sitúan en una media de 13 segundos y no presentan diferencias significativas entre pacientes y sujetos sanos. Respecto a los tiempos de fonación, tanto en /z/ como en /a/, los valores también son inferiores si los comparamos con los publicados por Wuyts *et al.* (2000), Gelfer y Pazera (2006) y Hakkesteegt *et al.* (2006). Aunque coincidimos con Wuyts *et al.* en la percepción de una alta variabilidad de dicho parámetro. Podemos destacar el acuerdo

con Rodríguez-Parra *et al.* (2009) en cuanto al TMF en /a/ del grupo experimental situado en los 9 segundos. Finalmente, nos cabe señalar que junto con Speyer *et al.* (2010), las diferencias entre los TMF entre los grupos control y experimental son significativas aunque dichos autores presentan una diferencia promedio de seis segundos frente a los tres que arrojan nuestros datos. La explicación a este conjunto de disimilitudes entre los datos de nuestra investigación debemos encontrarlos en las posibles diferencias de instrucciones dadas a los sujetos participantes. Es probable que unas instrucciones con mayor o menor énfasis y gestualidad puedan favorecer tiempos de emisión más o menos largos. La diversidad de resultados observados en las medidas de tiempos de fonación y espiración da lugar a cocientes s/z también variables. Si comparamos nuestros datos con los de Gelfer y Pazera (2006), observamos que, en lo relativo al grupo control, ambos valores son muy cercanos a 1. Pero nuestros resultados presentan una media en el grupo experimental de 1.2 que está por debajo de los 1.4 que informan Rodríguez-Parra *et al.* (2009). En nuestros cálculos referidos al Q s/z, hemos obtenido que la diferencia entre grupos es significativa pero el estadístico AUC ha resultado bajo. Estos datos junto a la no coincidencia respecto a la literatura anterior nos induce a no considerar este cociente como un indicador discriminante de normalidad o patología y nos hace recordar a los profesionales clínicos la necesidad de interpretar los parámetros de eficiencia respiratoria y fonorespiratoria conjuntamente con la observación cualitativa del modo y el tipo respiratorio y en relación al acto fonador y comunicativo.

En el apartado a los resultados relativos a las frecuencias podemos destacar la práctica coincidencia con los datos ofrecidos por Baken (1999) respecto a la F₀ espontánea de hombres y mujeres. Nuestros resultados nos impulsan a considerar que el mejor procedimiento para obtener éste dato es el cálculo informatizado de la media de una producción espontánea, evitando el uso de vocales sostenidas. En cuanto a la F₀ máxima existe un cierto acuerdo en determinar que en los hombres sanos este valor se sitúa alrededor de los 670 Hz. (Baken, 1999; Hakkestee *et al.* 2006). Dicho valor en las mujeres sanas no se presenta tan homogéneo. En nuestra investigación, la media se situó en los 827 Hz, cercanos a los de 943 Hz. de los trabajos de Hakkestee *et al.* (2006) pero a unos seis semitonos por debajo de los datos de Baken (1999) y Ma y Yiu (2006). Estas diferencias no menguan el valor de las notas de corte determinadas en la curva ROC y nos anima a definir el Do₄ en varones y el Mi₄ en mujeres, como los valores que permiten identificar una limitación patológica en el parámetro de la frecuencia máxima.

Respecto a la intensidad mínima, nuestros datos del grupo control y experimental coinciden parcialmente con los estudios de Ma y Yiu (2006) y de Hakkesteeg *et al.* (2006), lo que unido a los bajos valores del estadístico AUC nos llevan a no considerar este parámetro como discriminante de la presencia de alteración vocal. Esta misma consideración debemos asignarla al parámetro *jitter*, pues los trabajos antes citados o los de Rodríguez-Parra *et al.* (2009) sitúan las medias de normalidad por debajo de 1.0 mientras que nuestros resultados lo sitúan en 1.1. En el caso del grupo experimental los valores son más altos en los trabajos de Hakkesteeg *et al.* (2006) y de Rodríguez-Parra *et al.* (2009) que en los nuestros. La variabilidad y dudosa fiabilidad de los datos obtenidos con el cálculo del *jitter* se ponen de manifiesto en los trabajos de Maryn *et al.* (2009), a los que nos adherimos, en los que concluyen que no se trata de un buen instrumento para la valoración de la calidad vocal.

Finalmente, en referencia al *Dysphonia Severity Index*, observamos que los resultados de nuestro estudio muestran claras diferencias entre el grupo control y el experimental y el valor promedio del grupo control coincide con los datos que aportan los estudios de sus creadores (Wuyts *et al.* 2000) y de la mayoría de trabajos que has estudiado (Maryn, Corthals, Van Cauwenberge, Roy y De Bodt, 2010) su validez y fiabilidad (Hakkesteeg *et al.* 2006, 2008 y 2010). Respecto a los resultados referidos al grupo experimental, éstos presentan un rango muy amplio en nuestro estudio y en los diversos trabajos consultados se pueden observar resultados también dispares influidos en gran medida por la selección de la muestra.

En conjunto, nuestros resultados se muestran más próximos a la literatura cuando se refieren al grupo control. Este hecho puede encontrar su explicación en los criterios de inclusión de nuestro grupo experimental. La selección se realizó en el marco de una investigación que persigue comprender el papel de las características psicológicas en las alteraciones vocales y a tal efecto se decidió excluir las alteraciones de origen neurológico como las parálisis laríngeas o las distonías y bajo el mismo criterio no se incorporaron los pacientes neoplásicos o presbifónicos. También somos conscientes que algunos resultados podrían matizarse si hubiéramos conocido con certeza los diagnósticos lesionales, pues algunas lesiones determinan algunos de los valores estudiados. En la muestra tampoco estaban presentes pacientes en espera de intervención quirúrgica por lesiones laríngeas, pues fueron captados en los centros de atención logopédica concertados en los que, desgraciadamente, se realizan pocos

tratamientos pre quirúrgicos. Probablemente la inclusión de todas estas tipologías de pacientes habría aumentado los valores relativos a la alteración vocal.

Finalmente podemos concluir el conjunto de medidas vocales y respiratorias que hemos estudiado son útiles para la práctica clínica. Unos deben usarse con cautela como factores determinantes en la consideración de normalidad o patología pero, integrados en la mirada global del evaluador, pueden aportar información muy relevante en la comprensión de cada caso y pueden servir de referente de su evolución tras las actuaciones médicas o logopédicas. En el caso de la frecuencia máxima podemos considerar que nuestro objetivo de dotar a la comunidad profesional de un valor de corte se ha conseguido. Finalmente, en referencia al *Dysphonia Severity Index* consideramos que este trabajo es un primer estudio experimental en nuestro país, pero que la inclusión de un valor tan discutido como el *jitter* le resta robustez y recorrido futuro. Por este motivo, y sólo mientras no aparezcan nuevos estudios en los que se compare directamente la validez de este tipo de índices multiparamétricos, consideramos al DSI como un indicador útil en el diagnóstico clínico de la voz por su facilidad de cálculo y accesibilidad de los instrumentos necesarios para su obtención.

Referencias

- Arias, C. (1992). Examen funcional de la fonación de la exploración foniátrica. Bases para la aplicación en logopedia. *Revista de Logopedia, Foniatría y Audiología*, 12(2), 111-118.
- Aronson, A. E. (1980). *Clinical voice disorders : A interdisciplinary approach*. NY: Thieme-Stratton.
- Awan, S. N., y Roy, N. (2006). Toward the development of an objective index of dysphonia severity: A four-factor acoustic model. *Clinical linguistics & phonetics*, 20(1), 35-49.
- Awan, S. N., y Roy, N. (2009). Outcomes measurement in voice disorders: application of an acoustic index of dysphonia severity. *Journal of Speech, Language, and Hearing Research*, 52(2), 482-499.
- Awan, S. N., Roy, N., y Dromey, C. (2009). Estimating dysphonia severity in continuous speech: application of a multi-parameter spectral/cepstral model. *Clinical Linguistics & Phonetics*, 23(11), 825-841.
- Awan, S. N., Roy, N., Jette, M. E., Meltzner, G. S., y Hillman, R. E. (2010). Quantifying dysphonia severity using a spectral/cepstral-based acoustic index:

Comparisons with auditory-perceptual judgements from the CAPE-V. *Clinical Linguistics & Phonetics*, 24(9), 742-758.

Baken, R. J. (1999). *Clinical measurement of speech and voice*. San Diego: Singular Pub. Group.

Behrman, A. (2005). Common practices of voice therapists in the evaluation of patients. *Journal of Voice*, 19(3), 454-469.

Burgueño, M. J., García-Bastos, J. L., y González-Buitrago, J. M. (1995). Las curvas ROC en la evaluación de las pruebas diagnósticas. *Medicina Clínica*, 104(17), 661-670.

Carding, P. N., Wilson, J. A., MacKenzie, K., y Deary, I. J. (2009). Measuring voice outcomes: State of the science review. *The Journal of Laryngology and Otology*, 123(8), 823-829.

Deary, I. J., Webb, A., Mackenzie, K., Wilson, J. A., y Carding, P. N. (2004). Short, self-report voice symptom scales: psychometric characteristics of the voice handicap index-10 and the vocal performance questionnaire. *Otolaryngology-Head and Neck Surgery*, 131(3), 232-235.

Dejonckere, P. H., Bradley, P., Clemente, P., Cornut, G., Crevier-Buchman, L., Friedrich, G., et al. (2001). A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques - Guideline elaborated by the Committee on Phoniatics of the European Laryngological Society (ELS). *European Archives of Oto-Rhino-Laryngology*, 258(2), 77-82.

Gelfer, M. P., y Pazera, J. F. (2006). Maximum duration of sustained /s/ and /z/ and the s/z ratio with controlled intensity. *Journal of Voice*, 20(3), 369-379.

Hakkestegt, M. M., Brocaar, M. P., y Wieringa, M. H. (2010). The applicability of the dysphonia severity index and the voice handicap index in evaluating effects of voice therapy and phonosurgery. *Journal of Voice*, 24(2), 199-205.

Hakkestegt, M. M., Brocaar, M. P., Wieringa, M. H., y Feenstra, L. (2006). Influence of age and gender on the dysphonia severity index: A study of normative values. *Folia Phoniatica et Logopaedica*, 58(4), 264-273.

Hakkestegt, M. M., Wieringa, M. H., Brocaar, M. P., Mulder, P. G., y Feenstra, L. (2008). The interobserver and test-retest variability of the dysphonia severity index. *Folia Phoniatica et Logopaedica*, 60(2), 86-90.

Hirano, M. (1981). *Clinical examination of voice*. Wien ; New York: Springer-Verlag.

- Jacobson, B. H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., y Benninger, M. S. (1997). The voice handicap index (VHI): Development and validation. *American Journal of Speech-Language Pathology*, 6, 66-70.
- Ma, E., Robertson, J., Radford, C., Vagne, S., El-Halabi, R., y Yiu, E. M. (2007). Reliability of speaking and maximum voice range measures in screening for dysphonia. *Journal of Voice*, 21(4), 397-406.
- Ma, E., y Yiu, E. M. L. (2006). Multiparametric evaluation of dysphonic severity. *Journal of Voice*, 20(3), 380-390.
- Maryn, Y., Corthals, P., Van Cauwenberge, P., Roy, N., y De Bodt, M. (2010). Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels. *Journal of Voice*, 24(5), 540-555.
- Maryn, Y., De Bodt, M., y Roy, N. (2010). The acoustic voice quality index: Toward improved treatment outcomes assessment in voice disorders. *Journal of Communication Disorders*, 43(3), 161-174.
- Neiman, G. S., y Edeson, B. (1981). Procedural aspects of eliciting maximum phonation time. *Folia phoniatica et Logopaedica*, 33(5), 285-293.
- Núñez-Batalla, F., Corte-Santos, P., Senaris-Gonzalez, B., Llorente-Pendas, J. L., Górriz-Gil, C., y Suárez-Nieto, C. (2007). Adaptación y validación del índice de incapacidad vocal (VHI-30) y su versión abreviada (VHI-10) al español. *Acta Otorrinolaringologica Española*, 58(9), 386-392.
- Oates, J. (2009). Auditory-perceptual evaluation of disordered voice quality: pros, cons and future directions. *Folia phoniatica et logopaedica*, 61(1), 49-56.
- Rodríguez-Parra, M. J., Adrian, J. A., y Casado, J. C. (2009). Voice therapy used to test a basic protocol for multidimensional assessment of dysphonia. *Journal of Voice*, 23(3), 304-318.
- Rosen, C. A., Lee, A. S., Osborne, J., Zullo, T., y Murry, T. (2004). Development and validation of the voice handicap index-10. *The Laryngoscope*, 114(9), 1549-1556.
- Speyer, R., Bogaardt, H. C., Passos, V. L., Roodenburg, N. P., Zumach, A., Heijnen, M. A., et al. (2010). Maximum phonation time: Variability and reliability. *Journal of Voice*, 24(3), 281-284.
- Steen, I. N., MacKenzie, K., Carding, P. N., Webb, A., Deary, I. J., y Wilson, J. A. (2008). Optimising outcome assessment of voice interventions, II: Sensitivity to change of self-reported and observer-rated measures. *The Journal of Laryngology and Otology*, 122(1), 46-51.

- Verdonck-de Leeuw, I. M., Kuik, D. J., De Bodt, M., Guimaraes, I., Holmberg, E. B., Nawka, T., et al. (2008). Validation of the voice handicap index by assessing equivalence of European translations. *Folia Phoniatica et Logopaedica*, 60(4), 173-178.
- Webb, A., Carding, P. N., Deary, I. J., MacKenzie, K., Steen, N., y Wilson, J. A. (2004). The reliability of three perceptual evaluation scales for dysphonia. *European Archives of Oto-rhino-laryngology*, 261(8), 429-434.
- Webb, A., Carding, P. N., Deary, I. J., MacKenzie, K., Steen, I. N., y Wilson, J. A. (2007). Optimising outcome assessment of voice interventions, I: Reliability and validity of three self-reported scales. *The Journal of Laryngology and Otology*, 121(8), 763-767.
- Wuyts, F. L., De Bodt, M., Molenberghs, G., Remacle, M., Heylen, L., Millet, B., et al. (2000). The dysphonia severity index: an objective measure of vocal quality based on a multiparameter approach. *Journal of Speech, Language, and Hearing Research*, 43(3), 796-809.
- Zraick, R. I., Risner, B. Y., Smith-Olinde, L., Gregg, B. A., Johnson, F. L., y McWeeny, E. K. (2007). Patient versus partner perception of voice handicap. *Journal of Voice*, 21(4), 485-494.