UNIVERSIDAD DE EXTREMADURA

TESIS DOCTORAL

DESARROLLO Y APLICACIÓN DE UN PROGRAMA BASADO EN BIOMARCADORES ACÚSTICOS PARA LA PREVENCIÓN DE PATOLOGÍAS DE LA VOZ

MARÍA SANDRA PANIAGUA VIVAS

PROGRAMA DE DOCTORADO EN SALUD PÚBLICA Y ANIMAL

2019

UNIVERSIDAD DE EXTREMADURA

TESIS DOCTORAL

DESARROLLO Y APLICACIÓN DE UN PROGRAMA BASADO EN BIOMARCADORES ACÚSTICOS PARA LA PREVENCIÓN DE PATOLOGÍAS DE LA VOZ

MARÍA SANDRA PANIAGUA VIVAS

PROGRAMA DE DOCTORADO EN SALUD PÚBLICA Y ANIMAL

Conformidad de los directores

Fdo.: Carlos Javier Pérez Sánchez Fdo.: Fernando Calle Alonso

2019

Resumen

El uso de la voz como herramienta de trabajo supone para el profesional de la voz (docente, locutor, teleoperador…) un riesgo de padecer patologías vocales. Diversos organismos internacionales consideran al colectivo docente como el que presenta mayor riesgo de padecer trastornos vocales. Los nódulos vocales constituyen la única enfermedad de la voz reconocida legalmente como enfermedad profesional en España. Actualmente, los análisis acústicos basados en características extraídas automáticamente de grabaciones de voz empiezan a jugar un papel importante para el diagnóstico y seguimiento de estas patologías. El objetivo de esta tesis ha sido desarrollar un programa de prevención de trastornos vocales basado en características acústicas que permita detectar riesgos vocales para, posteriormente, abordarlos. Este programa se ha aplicado a docentes de la Universidad de Extremadura. A partir de características acústicas lineales y no lineales extraídas de grabaciones de voz, los profesores se han asignado a distintos grupos de riesgo de padecer nódulos vocales o trastornos con similares características de fonación. Según el grupo de riesgo asignado, se han propuesto medidas de actuación en diferentes niveles prevención (primaria, secundaria y terciaria), que conllevan acciones encaminadas a la protección, promoción, prevención y rehabilitación de la salud vocal. Los resultados obtenidos en el experimento muestran que es posible la utilización de este tipo de tecnologías no invasivas y de bajo coste como ayuda al diagnóstico clínico de patologías vocales. Aunque el programa se ha aplicado a profesores de universidad, se puede extender a cualquier otro colectivo de profesionales de la voz.

Palabras claves: Características acústicas; disfonía; docente; nódulos vocales; profesional de la voz; programa de prevención; sistema de ayuda al diagnóstico.

Abstract

The use of voice as a work tool implies a risk of vocal pathologies for the voice professionals (teacher, speaker, telemarketer...). Several international organizations consider teachers as the professionals with the highest risk of vocal disorders. Vocal fold nodules constitute the only voice disease legally recognized as a professional disease in Spain. Currently, acoustic analyses based on features automatically extracted from voice recordings begin to play an important role in the diagnosis and monitoring of these pathologies. The objective of this PhD has been to develop a voice care program for the prevention of vocal fold disorders based on acoustic features that allows the detection of vocal risks for subsequently addressing. This program has been applied to teachers of the University of Extremadura. Based on linear and non-linear acoustic features extracted from voice recordings, teachers have been assigned to different risk groups of having vocal nodules or disorders with similar phonation characteristics. According to the assigned risk group, several measures have been proposed at different levels of prevention (primary, secondary and tertiary), which involve actions aimed at the protection, promotion, prevention and rehabilitation of the vocal health. The results obtained in the experiment show that the use of this type of non-invasive and low-cost technologies is possible as a diagnostic support system for vocal pathologies. Although the program has been applied to university teachers, it can be extended to any other group of voice professionals.

Keywords: Acoustic features; dysphonia; teacher; vocal fold nodule; voice professional; voice prevention program; diagnostic support system.

A mis padres, ejemplo de lealtad y cariño.

A Juan, mi eterno norte...Me regalaste una estrella, te daría una vida.

A mis perritas, vuestros ojos han sido mil sonrisas.

Agradecimientos

A mi director de tesis, el Dr. Carlos Javier Pérez Sánchez, por su dedicación, su impecable trabajo, sus enseñanzas y su paciencia en momentos difíciles.

A mi codirector de tesis, el Dr. Fernando Calle Alonso, por su implicación, su trabajo, sus ánimos y su comprensión siempre. Sin él esto no hubiera podido empezar.

A la Dra. Carmen Salazar Cabrera, Jefa del Servicio de Otorrino del Hospital Universitario de Cáceres, por abrirme la puerta a su servicio y dejarme formar parte de su consulta.

A Esther y Elena, mis compañeras en la Unidad de Voz del Hospital Universitario de Cáceres, por considerarme una más en su equipo, y procurar que no se perdiera ningún paciente. Sin vuestra colaboración esto no hubiera sido posible. A Pepi, Ángela y las demás compañeras TCAE que siempre me ayudaron. Al resto del Servicio de Otorrino del Hospital Universitario de Cáceres por el respeto mostrado siempre por esta investigación.

Al Dr. Antonio J. Moreno Gómez, Jefe del Servicio de Prevención de la Universidad de Extremadura, por su asesoramiento e inclusión del experimento en el programa de prevención de riesgos laborales

A Gloria Grajales y Antonio Pérez, enfermer@s del Servicio de Prevención de la Universidad de Extremadura, por su ayuda en el proceso de recogida de grabaciones y encuestas.

A Mario Madruga, por su inestimable ayuda y participación en la extracción de las características acústicas de las grabaciones de voz.

A Maritina, por abrirme la primera puerta en una todavía Escuela Universitaria. Tu espíritu de lucha por la profesión que compartimos, de superación y de romántico inconformismo han sido una valiosa enseñanza para mí. Esta tesis es también un poquito tuya.

A mis amigos, por apostar siempre por mí en este camino. A mis compañer@s de dirección de Mérida, por alentarme, comprenderme y hacerme ver que siempre se puede.

A todos/as aquellos/as que participaron de forma voluntaria en esta investigación, especialmente a los pacientes, que aun teniendo dificultades fueron capaces de repetir cuatro veces seguidas la vocal /a/.

Este proyecto ha contado con el apoyo financiero de los proyectos MTM2014-56949-C3-3-R y MTM2017-86875-C3-2-R (Ministerio de Ciencia, Innovación y Universidades), y de los proyectos IB16054 y GR18108 (Junta de Extremadura/Fondo Europeo de Desarrollo Regional, Unión Europea).

Índice general

Resumen i
Abstract iii
Agradecimientos vii
Índice de figuras xi
Índice de tablas xiii
Capítulo 1. Introducción 1
Capítulo 2. Bases de la fonación 5
2.1. Antecedentes históricos de la fonación 5
2.2. Bases anatómicas de la fonación 7
2.1.1. El fuelle del aparato fonador 8
2.1.2. El vibrador del aparato fonador. La laringe 9
2.1.3. El resonador del aparato fonador 13
2.2. Bases fisiológicas de la fonación 15
2.2.1. Desarrollo de la laringe 15
2.2.2. Funciones de la laringe 15
2.2.3. El ciclo vocal y la onda mucosa 17
2.2.4. Mecanismo vibratorio de la cuerda vocal 19
2.3. Bases físico-acústicas de la fonación 20
2.3.1. Propiedades de las ondas sonoras. El sonido 20
2.3.2. Modelo acústico de generación de voz. Los formantes 21
2.4. Análisis acústico y estudio funcional de la voz 22
2.4.1. Parámetros de frecuencia 24
2.4.2. Parámetros de perturbación de amplitud 26
2.4.3. Parámetros de ruido 28
2.4.4. Parámetros acústicos no lineales 28
2.4.5. Estudio funcional de la voz 29
Capítulo 3. Programa de prevención de riesgos en la voz 33
3.1. La voz normal y la voz patológica 33
3.1.1. Disfonías Funcionales 35
3.1.2. Disfonías Orgánicas 36
3.1.3. Disfonías orgánico-funcionales 37
3.1.4. Características vibratorias de nódulos, pólipos y edema de Reinke 42

3.2. Uso profesional de la voz 44
3.3. Salud vocal del docente 45
3.3.1. Factores de riesgo vocal 46
3.3.2. Niveles de prevención 49
3.4. Programa de prevención de disfonías en docentes 50
3.4.1. Asignación de docentes a grupos de riesgo 50
3.4.2. Asignación de niveles de prevención 51
Capítulo 4. Materiales y métodos 57
4.1. Diseño del estudio 57
4.1.1. Descripción de la muestra 58
4.1.2. Criterios de inclusión y exclusión 58
4.1.3. Recogida de datos 59
4.2. Variables 61
4.3. Material técnico 64
4.4. Métodos estadísticos 65
Capítulo 5. Resultados experimentales 69
5.1. Análisis del cuestionario 69
5.2. Análisis de las características acústicas 74
5.3. Aplicación de modelos predictivos 80
5.4. Asignación de profesores a niveles de prevención 82
Capítulo 6. Discusión 89
Capítulo 7. Conclusión 101
Referencias 103

Índice de figuras

Figura 2.1. Retrato de Manuel García (Pérez de Urbina & Pérez, 2006). 7
Figura 2.2. Estructura anatómica de la Laringe (Dosal, 2014). 10
Figura 2.3. Morfología microscópica de los pliegues vocales (Giovanni, Lagier & Remacle 2010). 17
Figura 2.4. Fases del ciclo vibratorio (Calvache, 2015). 18
Figura 2.5. Parametrización de onda sonora (Goller, 2014). 21
Figura 2.6. Frecuencia fundamental y harmónicos (Goller, 2014). 24
Figura 2.7. Jitter y Shimmer en una señal acústica (Teixeira & Gonçalves, 2014). 26

Figura 3.1. Vista de nódulos vocales en cuerdas vocales en abducción (izquierda) y aducción (derecha) (Núñez & Fernández, 2013). 38
Figura 3.2. Visa de pólipos vocales (Núñez & Fernández, 2013). 39
Figura 3.3. Vista de pseudoquiste seroso (Núñez & Fernández, 2013) 40
Figura 3.4. Vista de edema de Reinke (Bonilla, Pérez & Camacho, 2018). 40
Figura 3.5. Vista de quiste subepitelial (Chuang, Badía & Napolitano, 2017). 41

Figura 4.1. Equipo de fibrolaringoscopio y estroboscopio 64
Figura 4.2. Micrófono Cardioide AKG-520 65
Figura 4.3. Tarjeta TASCAM US-322 65

Figura 5.1. Diagrama de cajas y bigotes para edad en los distintos grupos. 71
Figura 5.2. Diagrama de cajas y bigotes para el número de años fumando por grupos. 72
Figura 5.3. Diagrama de cajas y bigotes para las puntuaciones de VHI por grupos. 72
Figura 5.4. Histograma del tiempo de docencia en años del grupo de profesores. 73
Figura 5.5. Histograma del número de horas de clases semanales del grupo de profesores. 74
Figura 5.6. Diagrama de cajas y bigotes para Jitter relativo (izquierda) y Shimmer local (derecha) por grupos. 75
Figura 5.7. Diagrama de cajas y bigotes para GNE por grupos. 76
Figura 5.8. Diagrama de cajas y bigotes para DFA por grupos. 78
Figura 5.9. Diagrama de cajas y bigotes para RPDE (izquierda) y PPE (derecha) por grupos. 79
Figura 5. 10 Diagrama de cajas y bigotes para D2 (izquierda) y PermutEn (derecha) por grupos. 79
Figura 5.11. Diagrama de cajas y bigotes para FuzzyEn (izquierda) y SampleEn (derecha) por grupos. 79
Figura 5.12. Diagrama de cajas y bigotes de las variables SampleEn (arriba izquierda), D2 (arriba derecha), PPE (abajo izquierda), GNE (abajo derecha). 81
Figura 5.13. Diagrama de cajas y bigotes para VHI por grupos. 83
Figura 5.14. Diagrama de cajas y bigotes para VHI por nivel de prevención asignado. 87
Figura 5.14. Diagrama de cajas y bigotes para VHI por nivel de prevención asignado. 87

Índice de tablas

Tabla 5.1. Tabla de contingencia para las variables cualitativas del cuestionario. 69
Tabla 5.2. Estadísticos descriptivos para variables cuantitativas del cuestionario VHI. 70
Tabla 5.3. Tabla de descriptivos para las variables lineales. 75
Tabla 5.4. Tabla de descriptivos para las variables no lineales. 77
Tabla 5.5. Métricas estadísticas del rendimiento del clasificador con los predictores seleccionados. 80
Tabla 5.6. Estadísticos descriptivos y p-valores para las pruebas de comparación de medias de las características acústicas seleccionadas. 82
Tabla 5.7. Tabla de contingencia para la variable problemas de voz y grupo de riesgo asignado. 84
Tabla 5.8. Tabla de contingencia para la variable problemas de voz y los niveles de prevención asignados. 85
Tabla 5.9. Tabla de contingencia de los niveles de prevención en los distintos sexos . 86
Tabla 5.10. Tabla de contingencia de los niveles de prevención según condición de fumador 87

Capítulo 1. Introducción

Definía Platón la voz como “un impacto de aire que llega por los oídos al alma”. Son múltiples y dinámicos los fenómenos que participan en la generación de la voz humana. La voz, forma parte de la cultura de un pueblo, siendo el instrumento de comunicación por excelencia. Esta dimensión la ha convertido en protagonista de procesos sociales, personales y laborales. La relevancia de su uso y, por ende, la dificultad o imposibilidad del mismo cuando están presentes trastornos vocales, ha supuesto un creciente interés por su estudio a lo largo de los años.

Este trabajo ha centrado la atención en el uso profesional de la voz, concretamente en el que realizan profesionales de la enseñanza. La Organización Internacional del Trabajo (OIT) reconoce a los docentes como el primer colectivo profesional con riesgo de padecer enfermedades de la voz (Behlau et al., 2012; Farías, 2013; Niebudek-Bogusz & Śliwińska-Kowalska, 2013; Cantor et al., 2013). El uso de la voz de forma continuada sin entrenamiento supone un riesgo de carga vocal excesiva, que en ocasiones conlleva la aparición de lesiones vocales persistentes. Las demandas vocales a las que se enfrenta un profesor o maestro en el desarrollo de su actividad laboral diaria suponen una adaptación continua de su función fonatoria. Aspectos relacionados con componentes personales del individuo (estilos de vida, higiene vocal, factores emocionales y rasgos de personalidad, técnica vocal) y con ambientes y circunstancias de trabajo, inciden de forma directa e indirecta sobre la salud vocal. En España, hasta el año 2006 no estaba reconocida ninguna enfermedad de la voz como enfermedad profesional, es entonces cuando los nódulos vocales se incluyen en el cuadro de enfermedades profesionales con el código 2L0101 (RD 1299/2006 de 10 de noviembre) dentro del grupo “Enfermedades producidas por agentes físicos”, precisando como actividades capaces de producirlos aquellas “actividades en las que se precise un uso mantenido y continuo de la voz, como son profesores, cantantes, actores, teleoperadores, locutores…”.

La falta de aplicación de programas de prevención de salud vocal para colectivos en los que la voz es la principal herramienta de trabajo conlleva con frecuencia consecuencias negativas tanto en la salud de los profesionales, como en costes económicos para la administración y las empresas. El proceso patológico de la disfonía en el profesor requiere la puesta en marcha de acciones encaminadas a prevenir y tratar los factores de riesgo desencadenantes de patologías vocales, así como a crear las condiciones necesarias para que los comportamientos del profesor en el aula sean adecuados para la salud vocal.

Diferentes autores han estudiado diversos niveles de intervención preventiva en los que incluyen aspectos relacionados con los conocimientos del profesor sobre los procesos vocales patológicos, entrenamiento en técnica vocal, actuación sobre los ambientes de trabajo o educación sobre hábitos de higiene vocal y hábitos de vida saludables (Nanjundeswaran et al., 2012; Bovo et al., 2007; Sodersten et al., 2002; Castejón, 2014). Sin embargo, ninguno de los programas encontrados, relaciona el riesgo de padecer una patología vocal con el tipo de intervenciones a llevar a cabo.

En los últimos años se han desarrollado diferentes técnicas para la ayuda al diagnóstico y seguimiento de enfermedades relacionadas con la voz. El monográfico presentado por Baghai-Ravary y Beet (2013) proporciona una visión del estado del arte de la clasificación automática de señales de voz con propósitos clínicos. Esta línea de investigación no solo se circunscribe a enfermedades propias de otorrinolaringología, sino que se puede utilizar, por ejemplo, para la ayuda a la detección precoz y seguimiento de enfermedades neurodegenerativas que tienen efectos sobre la voz, como es el caso de la enfermedad de Parkinson (Little et al., 2009; Tsanas et al., 2011). Se ha considerado este enfoque para definir un programa preventivo con el que se demuestra que es posible asignar profesores a distintos grupos de riesgo de padecer una patología vocal, y asociarlos a diferentes niveles de atención preventiva.

El programa preventivo que se propone y se aplica con profesores universitarios, utiliza características acústicas extraídas automáticamente de grabaciones de voz. Mediante modelos de clasificación asigna la probabilidad de padecer nódulos vocales o trastornos compatibles con estos. Esta metodología es no invasiva y de bajo coste, y ayuda al diagnóstico clínico a través de parámetros objetivos.

Por tanto, el objetivo principal de esta tesis ha sido desarrollar un programa intervencionista de prevención de trastornos vocales basado en características acústicas que ha sido probado con profesores de la Universidad de Extremadura, y cuya finalidad es prevenir, retrasar o tratar nódulos vocales u otras patologías vocales con similares patrones de fonación.

Como objetivos parciales se han propuesto:

- Definir un programa preventivo de salud vocal con tres niveles de atención.
- Diseñar y aplicar un protocolo de grabación y encuesta para la recolección de datos de tres grupos diferentes de sujetos: profesores de la Universidad de Extremadura, pacientes con diagnóstico de nódulos vocales y sujetos con buena salud vocal.
- Analizar los resultados proporcionados por las encuestas y las características extraídas en los diferentes grupos considerados en un contexto comparativo.
- Aplicar los métodos de selección de características y clasificación para obtener la probabilidad de tener nódulos vocales o trastornos de la voz compatibles con ellos, para crear tres grupos de riesgo que se asignen a tres niveles de prevención.
- Valorar los resultados obtenidos.

El presente trabajo de tesis doctoral está estructurado en siete capítulos. En el Capítulo 1 se realiza una primera aproximación al tema, presentando la motivación para la realización del estudio presentado y los objetivos planteados.

El Capítulo 2 trata sobre las bases de la fonación. Comienza con una breve contextualización histórica sobre el desarrollo de los estudios de la fonación. A continuación, se describen las principales estructuras anatómicas que forman el aparato fonador y los principios fisiológicos que marcan su funcionamiento. También se explican cuáles son las bases físico-acústicas de la fonación, así como algunos parámetros acústicos que pueden analizarse durante la misma.

En el Capítulo 3 se define el programa de prevención. En una primera parte se explican los conceptos de voz normal y voz patológica, describiendo los distintos tipos de disfonías que se pueden encontrar. La segunda parte del capítulo está dedicada a describir aspectos relacionados con el uso profesional de la voz, la salud vocal de los docentes y los riesgos de desarrollo de patologías vocales a los que están expuestos. Por último, se presenta una propuesta de intervención preventiva para el colectivo de docentes de la Universidad de Extremadura, en la que se describen las actuaciones que podrían llevarse a cabo en los tres niveles de actuación definidos.

El Capítulo 4 describe los materiales y métodos utilizados en el experimento realizado. Recoge la información principal de las muestras utilizadas y los criterios de inclusión y exclusión considerados. Asimismo, explica cómo ha sido la recogida de datos y el material técnico empleado, las variables estudiadas, y los métodos estadísticos elegidos para el análisis.

Los resultados del experimento realizado se recogen en el Capítulo 5. Este capítulo analiza las variables cualitativas y cuantitativas más destacadas del cuestionario realizado, así como las características acústicas tenidas en cuenta en la investigación. También se muestra cómo ha sido la aplicación de los modelos predictivos y cuáles son sus resultados de asignación a grupos de riesgo, así como los niveles preventivos asociados.

En el Capítulo 6 se presenta la significación del estudio realizado en comparación con otras investigaciones relacionadas, resaltando la singularidad de la investigación en cuanto a la utilización de asociaciones de medidas de intervención preventiva basadas en los resultados de un modelo de clasificación fundamentado en características acústicas.

El Capítulo 7 está dedicado a presentar la conclusión de la tesis. Finalmente, se presentan las referencias bibliográficas.

Capítulo 2. Bases de la fonación

La voz humana es el resultado de la acción coordinada de numerosas estructuras corporales (órganos, aparatos y sistemas). Su evaluación y análisis requiere la integración y coordinación de numerosas técnicas que tienen como objetivo la evaluación de la calidad vocal.

Este capítulo está dividido en cuatro secciones, en las que se abordarán aspectos relacionados con las bases anatómicas de la fonación y su fisiología, así como sus bases físico-acústicas y su análisis acústico.

2.1. Antecedentes históricos de la fonación

A lo largo de la historia, el mecanismo de la fonación ha sido objeto de numerosos estudios que hoy en día permanecen. En el siglo II Galeno de Pérgamo en sus estudios sobre la fisiología del cuerpo humano, compara el órgano emisor de voz con una flauta, en la que la tráquea constituiría el cuerpo.

Durante el renacimiento, son varios los autores que muestran su interés por el estudio anatómico de la laringe humana. Durante el siglo XV destacan los trabajos de anatomía de Leonardo da Vinci que, en forma de dibujos dio a conocer al mundo. Un ejemplo de ellos son sus “Estudios anatómicos de la laringe y la pierna” expuestos en la *Royal Library* de *Windsor*. En el siglo XVI, destaca Andrés Vesalio (*Andries van Wesel*), al que autores como Reverón y Romero (2007) han considerado fundador de la anatomía moderna por la publicación de su trabajo “*De humani corporis fabrica*” en 1543.

En el siglo XVII, con Girolamo Fabricius d’Acquapendente, surge el interés por conocer las bases fisiológicas de los procesos laríngeos del habla, así como algunas características de ésta, y las cualidades en la voz cantada. Destacan sus obras “*De Laryngis Vocis Instrumento*”, y “*De Brotorum loquela*”. Fabricius matiza los estudios de su maestro Aristóteles sobre la producción de ondas sonoras, sosteniendo que no sólo la unión de dos cuerpos sólidos puede producirlas, sino también la de dos cuerpos blandos. Además, considera al aire como el elemento que genera el sonido y le da forma, siendo éste, una consecuencia de una alteración del estado del aire, ya sea por compresión o vibración (Fernández et al., 2006).

Según recogen estudios sobre Girolamo F. d´Acquapendente (García-Tapia,1996), este reconoce por primera vez la existencia de cuatro cartílagos como integrantes de la laringe, señalando los Aritenoides como estructuras no integrantes de esta, y definiéndolos como punto de apoyo que le confieren movilidad. Redefine al “cartílago innominado” con el nombre de “Cricoides”, debido a la semejanza de éste con una pieza de marfil anular que los turcos colocaban en su dedo pulgar para lanzar flechas. Elabora una completa descripción sobre los músculos internos de la laringe, a los que atribuye la misión de abrir y cerrar la glotis, donde para él, a partir de una aspiración violenta, se forma la voz.

A lo largo del siglo XVII son varios los autores que realizan estudios sobre las bases fisiológicas de la fonación y articulación del lenguaje. Destacan, Martin Mersenne con su "*Traité d'harmonie*", *Géraud* de *Cordemoy* con "*Discours physique de la parole*" o Bernard Lamy, que describe el funcionamiento de las cuerdas vocales en su trabajo "*La réthorique ou l'art de parler*" (Fernández, 2006).

Denis Dodart, a finales del siglo XVII describe el mecanismo de producción de la voz humana y su regulación tonal en su obra "*Mémoire sur les causes de la voix de l'homme et de ses différents tons. Mémoires de mathématique et de physique de l'Académie royale des sciences."* destacando como aspecto fundamental, que la emisión de frecuencias elevadas es debida al alargamiento de los pliegues vocales (García-Tapia & Fernández, 2013).

Antoine Ferrein (1741) es considerado por algunos autores como el iniciador de la fisiología experimental de la laringe. Cirujano y profesor de anatomía en Marsella y París, realizó estudios con laringes humanas y de animales, cuyos resultados plasmó en su obra "*Sur l'organe immédiat de la voix et de ses différenstons*". Descubrió, que, tras soplar a través de la laringe, el sonido desaparecía al tocar las estructuras vibrantes. Sin embargo, cuando comprimía parcialmente el segmento anterior o posterior de la glotis, acortando la longitud del repliegue vocal, la frecuencia de vibración ascendía produciendo un tono más agudo, semejante al que emite un instrumento musical de cuerda, al acortar una de ellas. Estas similitudes entre el órgano vocal y los instrumentos de cuerda, le llevaron a decir: "Esas bandas que denominaré por tanto cuerdas vocales, pueden ser comparadas a las cuerdas dobles del clavicordio". Esta denominación es la que ha permanecido vigente hasta nuestros días (García-Tapia, 1996).

Henri Dutrochet y Françoise Magendie, a principios del siglo XIX, destacan la participación del músculo Tiroaritenoideo en la modificación del tono de voz. Magendie estudia los cambios que se producen con la intervención de los nervios laríngeos, concluyendo que el cierre de la glotis depende del nervio laríngeo superior y que la apertura de la glotis depende del nervio laríngeo inferior (Hasson, 1962).

Johannes Müller, profesor de fisiología en Berlín, presenta en 1837 el resultado de los estudios realizados sobre laringes aisladas obtenidas de animales y cadáveres humanos (Gould et al., 1993). En ellos, refleja aspectos como que la voz es el resultado de la vibración de las cuerdas vocales al paso del aire, o la influencia de la tensión de las cuerdas vocales a la hora de emitir un tono u otro de voz. Estas conclusiones, serán la base años más tarde, para la formulación de distintas teorías sobre la fisiología de la fonación, como la Teoría Mioelástica de la fonación, que completó y difundió Janwillem van den Berg en 1958.

En 1854, un cantante de ópera y profesor de canto español, Manuel García (véase la Figura 2.1) es considerado la primera persona en visualizar en un sujeto vivo, la laringe y la porción superior de la tráquea mediante un instrumento dotado de dos espejos de dentista colocados en serie, y utilizando la luz solar, que recibirá el nombre de laringoscopio.

Manuel García da a conocer el laringoscopio, la técnica de uso y las primeras observaciones fisiológicas ante la *Royal Society of London* el 22 de marzo de 1855 con el título "*Physiological Observations on the Human Voice*" (Pérez de Urbina & Pérez, 2006).

Figura 2.1. Retrato de Manuel García (Pérez de Urbina & Pérez, 2006).

En 1962, Jorge Perelló, desarrolló la Teoría Muco-ondulatoria para explicar el comportamiento vibrador de las cuerdas vocales (García-Tapia, 1996). Años después, en 1975, Minoru Hirano describe por primera vez el concepto *body-cover complex* (cuerpo-cubierta) para referirse a la estructura de la cuerda vocal. Este concepto dará lugar al desarrollo de su Teoría Cuerpo-Cubierta para explicar la mecánica de la fonación (Hirano, 1989). Esta teoría es la que continúa vigente en la actualidad.

2.2. Bases anatómicas de la fonación

La generación, interpretación y respuesta del lenguaje están encuadradas en distintas áreas del cerebro (las áreas de Broca y *Wernicke*). La producción de la voz en el Sistema Nervioso Central no tiene una localización hemisférica concreta, sin embargo, al necesitar el *feed-back* auditivo, involucra a áreas temporales, así como pequeñas áreas frontales (memoria melódica: entonación, frecuencia, volumen).

La voz se produce gracias a la acción coordinada de distintas estructuras del cuerpo, que forman el sistema o aparato fonatorio. Para su estudio, el aparato fonador ha sido considerado como un solo órgano que puede dividirse en tres partes: el fuelle, el vibrador y el resonador. A continuación, se describe cada una de estas partes.

2.1.1. El fuelle del aparato fonador

El fuelle del aparato fonador está compuesto por estructuras del aparato respiratorio cuya ubicación anatómica las sitúa por debajo de las cuerdas vocales, y se denominan estructuras infraglóticas. Estas incluyen:

- Los pulmones como depósito de aire, la tráquea y los bronquios.
- Los músculos intercostales y abdominales esenciales en la espiración.
- El diafragma, como músculo inspirador.
- Otras estructuras complementarias: la caja torácica (costillas, esternón y columna vertebral (dorsal) y los músculos accesorios de la respiración (escalenos, esternocleidomastoideo y trapecios).

Cada pulmón, está recubierto por la pleura, una membrana serosa que tiene varias capas: la más externa (pleura costal), que se adhiere a la caja torácica y a la cara craneal del diafragma, y una más interna (pleura pulmonar), que recubre los pulmones. De esta forma, cualquier movimiento que realicen las costillas y el diafragma, será seguido por los pulmones.

El diafragma es un músculo circular abovedado que se encuentra dividiendo la cavidad torácica de la abdominal. Es el músculo de la respiración por excelencia (inspirador). Su inervación corre a cargo del nervio frénico, que se origina en la zona cervical. Consta de las siguientes partes:

- Porción anterior o porción esternal: se inserta en la cara posterior del apéndice Xifoides.
- Porción costal: costillas siete a doce.
- Porción lumbar o pilares del diafragma: se insertan en las tres primeras vértebras lumbares y tienen importantes relaciones anatómicas con el músculo Psoas y Cuadrado Lumbar.

Un patrón respiratorio coordinado es importante para conseguir una buena emisión de voz. Una adecuada relación fono-respiratoria marcará la diferencia en la calidad de la voz emitida. Durante la inspiración, el diafragma se contrae, aplanándose, mientras las costillas realizan dos tipos de movimientos: elevación en las más superiores, y lateralización en las más inferiores. A su vez, la musculatura abdominal se relaja, haciendo que se ensanche el perímetro abdominal. El descenso y contracción del diafragma produce un descenso de las vísceras abdominales, que empujan a la musculatura abdominal. Si ésta no está relajada, impedirá un correcto movimiento coordinado de todas las estructuras, produciéndose un freno en el descenso del diafragma, y una respiración inadecuada. Los músculos abdominales generan presión intraabdominal en determinados actos, como son la micción, la tos, el vómito o el parto. Al realizar una inspiración profunda, el aire queda atrapado entre los pliegues vocales (que se cierran) y el diafragma.

La respiración clavicular y torácica utiliza músculos del cuello como el Esternocleidomastoideo, y del tórax (intercostales), para producir una adecuada presión inspiratoria y espiratoria. Estos músculos, al contraerse crean tensiones que afectan y, en ocasiones, dificultan la fonación. La columna de aire generada al inspirar es empujada hacia arriba cuando los músculos abdominales se contraen, generando un movimiento ascendente de las vísceras abdominales. Esto a su a su vez, provoca que el diafragma empuje hacia arriba a los pulmones, y aumente la presión bajo los pliegues vocales (presión subglótica o infraglótica), que permanecerán cerrados hasta que la presión sea tan elevada que les obligue a abrirse y espirar con fuerza el aire. En el momento de la espiración pasiva (sin fonación), los pliegues vocales están abducidos, por lo que el aire pasa a través de ellos sin ningún esfuerzo, convirtiéndose en un proceso automático. Sin embargo, durante la espiración activa (fonación) los músculos abdominales se contraen, haciendo que el diafragma relajado suba, y los pliegues vocales se pongan en contacto, produciendo ahora una vibración controlada.

El sonido al hablar debe comenzar en el momento que se inicia la espiración, y las cuerdas vocales sufren una aproximación anatómica, constituyéndose así el comienzo del ciclo vocal. Para hablar o cantar, utilizaremos la respiración costo-diafragmática, también denominada respiración abdominal. Se produce en la parte más baja del tórax, y la más alta del abdomen, siendo el lugar donde radica el mayor control voluntario de la respiración. Durante el canto, el control voluntario de la espiración se convierte en eje central del acto. Las notas agudas necesitarán una mayor presión abdominal, de forma que los pliegues vocales vibren rápido y muy próximos, mientras que, en las notas graves, no es necesaria tanta fuerza de presión, ya que los pliegues se aproximan con menor fuerza, y vibran con menos velocidad. De la espiración también va a depender la intensidad o volumen de voz, de manera que, a mayor presión de los músculos abdominales, mayor amplitud de vibración de los pliegues vocales, y por lo tanto mayor intensidad de la voz resultante. Este equilibrio continuo entre el diafragma y los músculos abdominales será el que condicione una correcta presión de aire que facilite la fonación.

2.1.2. El vibrador del aparato fonador. La laringe

La laringe es un órgano impar que se sitúa en la región anterior del cuello. Su proyección sobre la columna cervical dependerá del sexo y de la edad. En varones adultos se sitúa entre las vértebras C3 y C7, mientras que en las mujeres suele estar un centímetro más alto, coincidiendo su borde inferior con la sexta vértebra cervical. En los niños su posición es aún más elevada, de manera que el borde libre de la epiglotis coincide con el borde superior de la tercera vértebra cervical. Tiene una forma tubular, con volumen y dimensiones condicionados también por la edad y el sexo. A partir de la vibración producida en las cuerdas vocales alojadas en la laringe, la energía continua del aire es transformada en energía discontinua: la voz. Atendiendo a su ubicación anatómica, la laringe puede dividirse en tres secciones. En dirección céfalo-caudal encontramos:

1. Supraglotis. Conforma la parte superior de la laringe y se sitúa entre el borde libre de la epiglotis y la cara superior de las cuerdas vocales. Se pueden distinguir distintas estructuras:

- Vestíbulo laríngeo
- Epiglotis
- Bandas ventriculares, conocidas como “cuerdas vocales falsas o superiores”
- Ventrículo de *Morgagni*

2. Glotis. Es un espacio triangular limitado por las cuerdas vocales (comisura anterior en la zona de unión de las mismas), las apófisis vocales y la cara interna de la base de los aritenoides (comisura posterior).
3. Subglotis. Es la parte inferior de la laringe, situada en el espacio entre la cara inferior de las cuerdas vocales y la parte más distal del cartílago cricoides, donde conecta con la tráquea.

La laringe está formada por un esqueleto de músculos, ligamentos y estructuras cartilaginosas que se articulan entre sí (véase la Figura 2.2).

Su estructura básica está compuesta de: esqueleto cartilaginoso y articulaciones que lo conectan; membranas y ligamentos; y músculos laríngeos intrínsecos y extrínsecos. A continuación, se describen en mayor profundidad estas partes de la estructura básica.

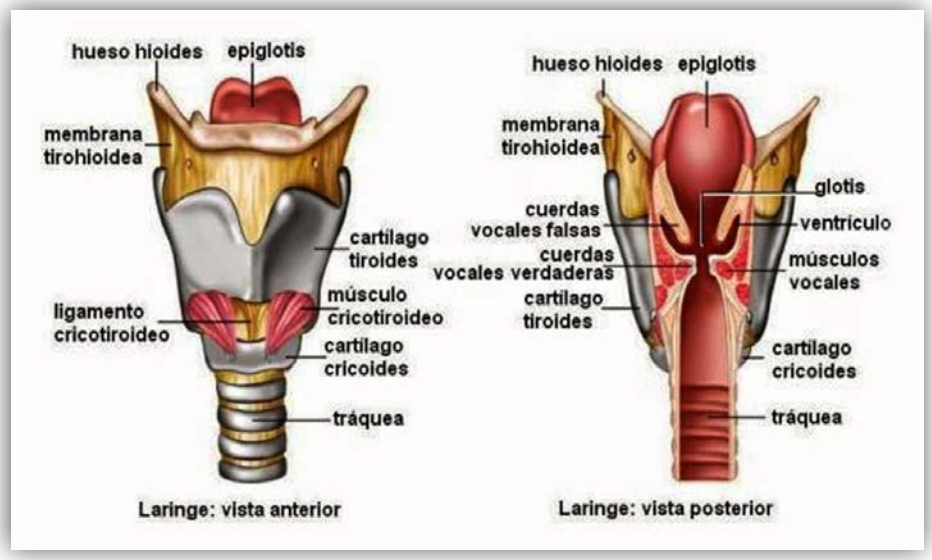

Figura 2.2. Estructura anatómica de la Laringe (Dosal, 2014).

El esqueleto cartilaginoso de la laringe está formado por una serie de cartílagos pares: Aritenoides, Corniculados o de *Santorini,* Cuneiformes o de *Morgagni* (o de *Wrisberg*) y Sesamoideos anteriores; y por tres cartílagos impares: Tiroides, Cricoides y Epiglótico.

Los Aritenoides son dos pequeños cartílagos hialinos bilaterales y simétricos, dispuestos a ambos lados de la línea media. Miden aproximadamente quince milímetros, y se sitúan por encima de la lámina del cricoides sobre la que se apoyan y con la que se articulan.

En ellos se diferencian dos apófisis bien definidas, una posterior y externa llamada proceso muscular, donde se insertan los músculos Cricoaritenoideo posterior y lateral, y otra anterior llamada apófisis vocal sobre la que se fija el extremo posterior del ligamento vocal. Los cartílagos Aritenoides tienen forma piramidal, y en cada uno de ellos se pueden encontrar una base, un vértice y tres caras:

- La cara interna es triangular, casi vertical y estrecha. Está recubierta por la mucosa de la escotadura Interaritenoidea o glotis Intercartilaginosa. Se retrae progresivamente de abajo a arriba, estrechándose por arriba, de forma que a partir de su mitad superior es tan solo un borde más que una autentica cara, en el que se inserta el ligamento Ariteno-epiglótico.
- La cara antero-externa es convexa en su conjunto. Destacan en ella dos depresiones separadas entre sí por la cresta arqueada de *Luschka* (*crista arcuata*): la Fóvea Triangular de localización superior y lugar de inserción del ligamento de la banda ventricular, y la Fóvea Oblonga o Fosita Oval, inferior a la cresta, donde se inserta el músculo Tiroaritenoideo.
- La cara posterior es excavada y lisa, y es el sitio de inserción del músculo Interaritenoideo

Los cartílagos Aritenoides al ser dos piezas móviles, permiten los movimientos de aproximación y separación de las cuerdas vocales, así como la regulación fina en la producción de la voz.

Los cartílagos Corniculados o de *Santorini* son dos estructuras cartilaginosas fibroelásticas, situadas inmediatamente por encima de los Aritenoides, reposando en su vértice. Tienen forma cónica, ligeramente aplastados de arriba abajo e incurvados de atrás hacia dentro. Su estructura está condicionada por la edad, de manera que a lo largo de los años se van convirtiendo en una prolongación hacia arriba de los Aritenoides, fusionándose con ellos, y formando un *crochet* cartilaginoso denominado *capitulum caritenoideum.*

Los cartílagos Cuneiformes o de *Morgagni,* también llamados cartílagos de *Wrisberg*, tienen una superficie aproximada de ocho milímetros, son cilíndricos y están situados por delante y por fuera de los Aritenoides y de los cartílagos de *Santorini*. Se encuentran formando parte del repliegue ari-epiglótico al que confieren rigidez, y en su extremidad superior forman el tubérculo de *Morgagni.*

Los cartílagos Sesamoideos son dos pequeñas formaciones ovoides cartilaginosas anteriores y posteriores, situadas en la escotadura del cartílago Tiroides.

El cartílago Tiroides está situado en la línea media, debajo del hueso Hioides y por encima del arco del cartílago Cricoides. Constituye junto con el cartílago Cricoides el armazón sobre el que se construye la forma de la laringe. Está formado por dos láminas trapezoidales o placas laterales (derecha e izquierda) que forman la mayor parte de la pared anterior y lateral de la laringe, y envuelve parcialmente a los demás cartílagos. Estas láminas son más anchas en el borde superior que en el inferior, y se unen por delante en la línea media con un ángulo (tiroideo o entrante) de 90º en el caso del hombre, y de 120º en el de la mujer.

La unión entre ambas láminas es incompleta en la zona supero-interna, dando lugar a la escotadura Tiroidea. La zona de unión de las láminas tiroideas da lugar al istmo del cartílago Tiroides, y se encuentra situado debajo de la escotadura Tiroidea. Orientado hacia abajo y atrás con un ángulo de 60º respecto al plano horizontal, da forma a una prominencia llamada tubérculo Tiroideo, más marcado y voluminoso en el hombre que en la mujer, y coloquialmente conocido con el término de "nuez de Adán".

El cartílago Cricoides constituye la base o pedestal de la laringe, por lo que también se le ha denominado cartílago Basal de *Ludwig*. Formado por dos segmentos de diferente altura, la lámina cricoidea y el arco cricoideo, su aspecto le confiere forma de anillo de sello, siendo más ancho por la parte de atrás, donde su orificio inferior, casi circular, alcanza un diámetro de 18 a 20 milímetros. Por la parte delantera, su diámetro anteroposterior, varía de los 20 a los 25 milímetros. Su orificio superior es oval, siendo su eje mayor oblicuo de arriba hacia abajo y de atrás hacia delante. A nivel de la lámina Cricoidea se encuentra la Cresta Media, en cuyos lados se sitúan las carillas articulares superiores, base del cartílago Aritenoides. En la unión de la lámina cricoidea y el arco Cricoideo se sitúan las carillas articulares inferiores, base para la unión con los cuernos inferiores del cartílago Tiroides. El cartílago Cricoides está unido al primer anillo traqueal a través de la membrana Cricotraqueal, siendo una referencia importante en la traqueotomía la depresión Intercricotiroidea del arco Cricoideo.

El Cartílago Epiglótico es una lámina delgada de tejido fibroelástico situado en la parte anterosuperior de la laringe. Está unido por su extremo inferior a la cara interna del ángulo del cartílago Tiroides mediante el ligamento Tiroepiglótico, y por la parte inferior de la cara anterior al hueso Hioides, a través del ligamento Hioepiglótico. Su función es poco importante a la hora de la fonación, pero sí lo es en la deglución, ya que se levanta durante la respiración, y desciende durante la deglución, permitiendo el paso del alimento al esófago, e impidiendo que pase a la glotis. A pesar de que la forma del cartílago Epiglótico asemeja a la hoja de una planta, hay ocasiones en las que su anatomía presenta una marcada concavidad transversal, denominándose "Epiglotis en Omega". Esta presentación, característica en laringes infantiles, dificulta la valoración del plano glótico en las exploraciones laríngeas.

Las membranas y ligamentos laríngeos se clasifican atendiendo a las superficies que unan. De manera que hablaremos de membranas y ligamentos intrínsecos al referirnos a aquellos que unen los cartílagos de la laringe entre sí, y extrínsecos, a aquellos que los conectan con estructuras próximas. Las membranas y ligamentos intrínsecos serían: ligamento Tiroepiglótico, ligamentos Aritenoepiglóticos, membrana Cricotiroidea y ligamentos Tiroaritenoides. Las estructuras englobadas dentro de las membranas y ligamentos extrínsecos son: membrana Tirohioidea, ligamentos Glosoepiglóticos, ligamento o membrana Hioepiglótica y membrana Cricotraqueal (Ramírez & Marfil, 2014)

Los músculos de la laringe son los responsables de los distintos movimientos llevados a cabo en ella, teniendo un papel relevante en los parámetros acústicos de la voz. Atendiendo a su función podemos hablar de:

- Músculos dilatadores: Cricoaritenoideo posterior.
- Músculos tensores: Cricotiroideo y Tiroaritenoideo
- Músculos constrictores: Cricoaritenoideo lateral, Tiroaritenoideo e Interaritenoideo

A su vez, podemos decir que la musculatura laríngea se divide en músculos extrínsecos y músculos intrínsecos. Los músculos extrínsecos son aquellos que se relacionan con los movimientos y fijación de la laringe, uniendo ésta con estructuras próximas, especialmente el hueso Hioides y el Esternón. Esto es así, gracias a que poseen una inserción en la laringe, y otra fuera de ella. Pueden dividirse en:

- Suprahioideos (grupo elevador de la laringe): Digástrico, Genihioideo, Milohiodeo y Estilohiodeo.
- Infrahiodeos (grupo depresor de la laringe): Omohioideo, Esternotiroideo, Tirohiodeo y Esternohioideo.

Los músculos propios de la laringe, o músculos intrínsecos, son aquellos cuyas dos inserciones están en la laringe, siendo los responsables del movimiento de las cuerdas vocales y de unir los cartílagos laríngeos entre sí. Esta musculatura actúa separando o acercando los pliegues vocales, produciendo la abertura o cierre comprendido entre ellos. Se habla, por tanto, de músculos abductores (músculo Cricoaritenoideo posterior) y músculos aductores (músculos Tiroaritenoideos, músculo Cricotiroideo, músculo Cricoaritenoideo lateral, músculo Aritenoideo transverso, músculo Aritenoideo oblicuo) respectivamente.

La inervación de la laringe proviene de dos porciones del X par craneal, o nervio Vago: el nervio laríngeo Superior y el nervio laríngeo Recurrente (o laríngeo Inferior). El nervio laríngeo superior se divide en una rama externa (función motora) que desciende para inervar al músculo Cricotiroideo o tensor de las cuerdas vocales, y una rama interna (función sensitiva) que penetra en la membrana Tirohioidea. Esta rama, a su vez se divide en una porción superior y una inferior, que inervan a las mucosas de la base de la lengua, la epiglotis y la laringe. Las lesiones de la rama externa del nervio laríngeo superior producen fatigas vocales y cambios en el timbre de la voz. La rama interna del laríngeo Superior (función sensitiva) da sensibilidad a la porción laríngea por encima del plano glótico, siendo su lesión la causante de anestesia laríngea y aspiración. El nervio laríngeo Recurrente inerva los músculos intrínsecos de la laringe. Se separa del vago a diferente nivel en cada lado, de manera que trabajan de forma independiente, posibilitando que la lesión en uno de los lados no produzca trastornos manifiestos en el otro.

2.1.3. El resonador del aparato fonador

El sonido producido en los pliegues vocales es modificado y amplificado en las distintas cavidades de resonancia.

Las estructuras resonadoras o resonadores son aquellas cavidades que se sitúan por encima de las cuerdas vocales, también denominadas cavidades supraglóticas. Estas cavidades u órganos de la articulación y resonancia están divididas en:

- Órganos resonadores: fosas nasales, cavidad oral y faringe.
- Órganos de la articulación: lengua, dientes, paladar, labios.

Para que la voz adquiera sonoridad es necesario que las cavidades de resonancia ajusten sus dimensiones a la tensión o frecuencia de vibración glótica. De la forma y el volumen de estas cavidades, dependerá el timbre de nuestra voz.

La boca es el principal resonador móvil, convirtiéndose en elemento fundamental en la voz hablada, cantada y proyectada. A partir de movimientos de la mandíbula, los labios o la lengua, se puede conseguir que la amplificación del sonido sea mayor o menor. En este sentido, la relajación o elevación del paladar blando, puede constituir que el sonido pueda pasar a la cavidad nasal o resonar exclusivamente en la boca, influyendo de forma definitiva en la calidad del timbre de la voz.

La musculatura de la cara juega un papel importante a la hora de conseguir un sonido u otro. Cuando está relajada, permite una buena función vocal, favoreciendo la apertura de la boca en favor de la gravedad, aplanando la lengua y subiendo el velo del paladar. En este sentido, destaca la función de los músculos implicados en la posición de los labios. El músculo orbicular permite colocar los labios en forma circular, y el buccinador favorece la abertura horizontal de la boca, llevando las comisuras de los labios hacia fuera.

El tamaño de la faringe provocará que el aire que circula por ella resuene con mayor o menor intensidad. La faringe en sus distintas porciones forma parte de los aparatos respiratorio y digestivo, estando implicada en los movimientos de deglución y respiración. Atendiendo a su localización se pueden distinguir distintas partes:

- Porción nasal (rinofaringe, nasofaringe o epifaringe).
- Porción oral (orofaringe, bucofaringe o mesofaringe).
- Porción laríngea (hipofaringe o laringofaringe).

Las lesiones o malformaciones de la faringe, en cualquiera de las localizaciones anteriores, interferirán en la cantidad del aire que circula por ellas, por tanto, en la calidad y timbre de la voz.

La cavidad nasal es un resonador fijo, ya que no podemos influir en su forma ni tamaño. Para utilizarlo, es necesario que el velo del paladar esté relajado, y el aire que circula a través de la faringe y llega a la boca, pueda conducirse hasta ella.

Es un resonador principalmente utilizado en el canto, y al igual que en el caso de la faringe, lesiones de cualquiera de sus estructuras, pueden establecer modificaciones en la resonancia de la voz.

2.2. Bases fisiológicas de la fonación

La capacidad del ser humano para producir sonidos articulados, que pudieran ser entendidos por otros iguales, es una característica que ha ido evolucionando a lo largo del tiempo.

2.2.1. Desarrollo de la laringe

El desarrollo de la laringe puede dividirse en etapa prenatal y postnatal. En la primera de ellas, durante el desarrollo embrionario, la laringe tiene una primera etapa de organogénesis y otra de maduración del órgano formado. Durante la 3ª semana del desarrollo embrionario (días 20-22), se empieza a formar en el embrión (2-3,5 milímetros), un surco medio, longitudinal, por debajo de las últimas bolsas branquiales, que se denomina surco o divertículo laringotraqueal. Esta estructura, se convertirá en un tubo o canal, denominado divertículo Respiratorio o Pulmonar, que derivará en lo que finalmente será el tracto respiratorio. La parte más craneal de este tubo dará lugar a la laringe y su extremo caudal se dividirá en dos esbozos bronquiales. Las cuerdas vocales, aparecerán como una diferenciación en la pared lateral de la laringe aproximadamente a los 40 días de gestación.

En el momento del nacimiento, la laringe se localiza en la parte alta del cuello, entre las vértebras C1 y C4, permitiendo movimientos simultáneos de respiración y deglución. Posteriormente, la laringe descenderá inferiormente hacia los dos años de edad, alcanzando una posición más inferior aún sobre los seis años, que la situará entre las vértebras C4 y C7. Esta nueva posición, supondrá un mejor rango de fonación, gracias a la mayor amplitud supraglótica de la faringe. Durante la pubertad, los cambios hormonales producen un crecimiento de la laringe en longitud y diámetro. Es en este momento cuando se produce la muda vocal, que en el sexo masculino será más evidente, ya que el crecimiento laríngeo oscila entre los 4 y los 11 milímetros, mientras que en el femenino oscila entre los 1,5 y los 4 milímetros (Torres, 2007). Además de estos cambios, también existen cambios histológicos en los pliegues vocales, resultando voces femeninas con hasta 2,5 semitonos de descenso, y voces masculinas con descenso de alrededor de una octava (Torres, 2007).

2.2.2. Funciones de la laringe

La laringe es una estructura móvil integrada en el aparato respiratorio, que cumple función protectora, respiratoria y fonatoria. La función protectora hace que la laringe actúe como una válvula o esfínter, que impide el paso de cualquier elemento que no sea aire a la vía aérea. Este mecanismo reflejo, se puede ver afectado por estímulos muy intensos, o manipulación directa. Esta capacidad valvular de la laringe permite alcanzar aumentos de volumen en la presión intratorácica durante la tos, el parto o la defecación.

La función respiratoria otorga a la laringe una participación activa en la regulación del flujo de aire entrante y saliente de los pulmones. En la inspiración, la glotis favorece el flujo de aire entrante, adoptando una forma alargada. Su retracción, permite un control y disminución de la velocidad del aire espirado, aumentando la presión intra-alveolar, y facilitando de esta forma el intercambio gaseoso. Durante la inspiración de la respiración no forzada, existe una relajación de los músculos constrictores y una contracción o tensión de los dilatadores, especialmente del Crico-aritenoideo posterior, máximo responsable de la dilatación glótica. Esto supone que la parálisis de los constrictores puede producir una dificultad respiratoria importante, pudiendo llegar a la asfixia por incompetencia muscular. El movimiento de las cuerdas vocales durante la respiración no forzada es mínimo, suponiendo la abducción de estas. El tamaño es un factor importante a tener en cuenta en la valoración de la función respiratoria de la laringe, ya que varía de unos individuos a otros, y afecta de manera directa en la resistencia mecánica que la laringe ofrece al aire.

La función fonatoria de la laringe incluye acciones mentales y físicas. La acción conjunta de la vibración de las cuerdas vocales más la amplificación y articulación en las cavidades resonadoras confieren a la voz las características y cualidades acústicas. A cada lado de la superficie interna de la laringe se sitúan dos pliegues superpuestos de su mucosa: los pliegues vestibulares (o ventriculares o bandas ventriculares) y los pliegues vocales (cuerdas vocales). Los pliegues vocales o cuerdas vocales son unas bandas blancas que se sitúan a ambos lados de la superficie interna de la laringe. Su función en la fonación es determinante, ya que se encargan de transformar la energía del aire, en energía acústica a partir de ondas sonoras. Hirano (1974) describió cinco capas de estructura histológica en el pliegue vocal (véase la Figura 2.3):

- Epitelio. Capa más superficial que recubre la lámina propia. Se trata de un epitelio delgado y lubrificado compuesto por células escamosas. Mantiene la forma de la cuerda vocal desde un punto de vista mecánico.
- Lámina superficial o espacio de Reinke. Constituida por tejido conjuntivo laxo, que permite la movilidad a la mucosa durante la fonación.
- Capa intermedia de la lámina propia. Constituida por fibras elásticas.
- Capa profunda de la lámina propia: constituida por fibras colágenas, forma junto con la capa intermedia el Ligamento vocal.
- Músculo vocal. Porción más profunda del pliegue vocal.

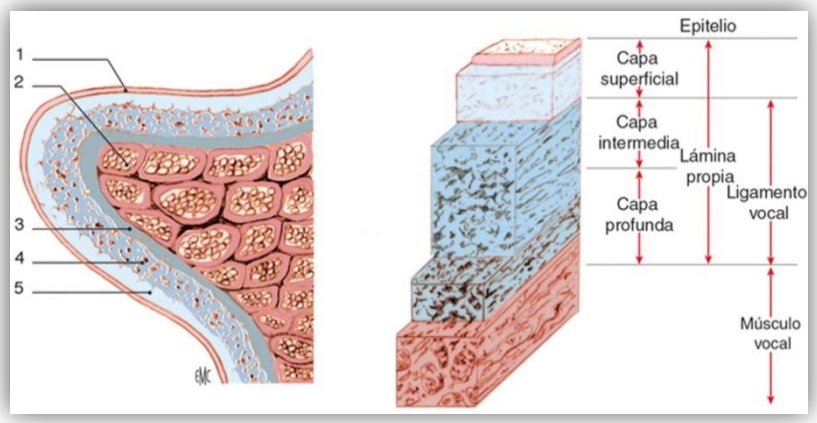

Figura 2.3. Morfología microscópica de los pliegues vocales (Giovanni, Lagier & Remacle 2010).

Para entender la mecánica de la fonación, las cinco capas de los pliegues vocales se agrupan en tres. La primera incluye la cubierta formada por epitelio y la capa superficial de la lámina propia. La segunda está formada por el ligamento vocal, o capa de transición, formado por la capa intermedia y profunda de la lámina propia. Y la tercera es el cuerpo, constituido por el músculo Vocal. Los músculos laríngeos controlan de forma mecánica la cubierta y la capa de transición. La contracción del Tiroaritenoideo controla de forma activa el cuerpo de los pliegues vocales, y el resto de músculos lo hacen de forma pasiva. Para determinar su funcionalidad, la cuerda vocal puede estudiarse en función de su histología, quedando dividida en una porción membranosa que ocupa los dos tercios anteriores, y donde se produce el fenómeno vibratorio en condiciones normales y de no patología, y una porción cartilaginosa que ocupa el tercio posterior.

2.2.3. El ciclo vocal y la onda mucosa

Se denomina ciclo fonatorio o ciclo vibratorio a cada una de las fases de abertura y cierre de los pliegues vocales (Torres, 2013), y fue descrito por primera vez por Schonhärl en 1960. En la Figura 2.4 se representan gráficamente las fases del ciclo fonatorio.

Para que se produzca el acto fonatorio, las cuerdas vocales deben estar en aducción, hecho que provoca la creación de un espacio supra y subglótico. Se pueden distinguir dos fases: una de transición o preparación, en la que existe una posición de relativa abertura de los pliegues vocales, y otra de ejecución o ataque, en la que se produce una oclusión casi completa (García-Tapia, 1996). Durante el proceso de aproximación de los pliegues vocales en la primera fase del ciclo vibratorio, se producen varios sucesos: reducción progresiva del flujo aéreo, aumento de la presión subglótica y aumento de la velocidad del flujo. Antes de la fonación, los pliegues vocales deben estar en contacto, de manera que la hendidura glótica permanezca cerrada. El flujo de aire generado en la espiración produce un aumento progresivo de la presión subglótica (o infraglótica).

Cuando la presión subglótica supera la presión de cierre de los pliegues vocales, éstos se separan, desde el borde más inferior hasta el más superior, permitiendo la salida del aire, de manera que la presión ejercida en la hendidura glótica descenderá bruscamente. El efecto *Bernoulli***,** a partir de la succión aplicada sobre los pliegues vocales, produce nuevamente la aproximación de la mucosa hacia la línea media. La fuerza del flujo de aire comienza a disminuir tras el cierre de los bordes inferiores de las cuerdas vocales, hasta el cierre de los superiores, de manera que la presión subglótica se vuelve a formar, y el ciclo comienza de nuevo. El principio de *Bernoulli*, utilizado para describir el comportamiento de fluidos al producirse un descenso de la presión del líquido en regiones donde la velocidad del flujo es mayor por el estrechamiento de la vía de paso, puede aplicarse al acto fonatorio, ya que la reducción de la presión lateral del aire, debido

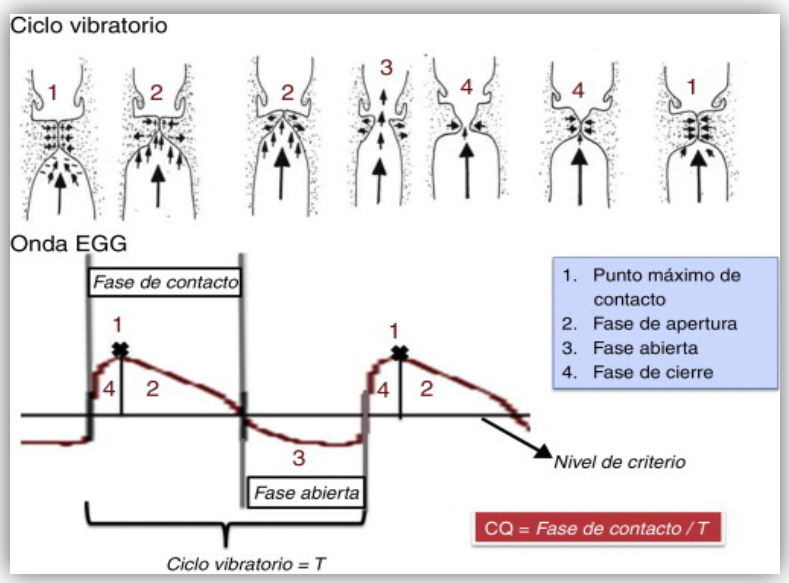

Figura 2.4. Fases del ciclo vibratorio (Calvache, 2015).

a su movimiento hacia delante, hace que la mucosa de los pliegues vocales se aproxime.

El fenómeno vibratorio interpretado como sonido, se produce, por tanto, por la salida de flujos de aire intermitentes. Cada escape de aire produce una onda de presión que se transmite hasta la cavidad faringo-bucal. La frecuencia del sonido emitido, denominada tono, será determinada por la cadencia de los flujos de aire que salen.

El término onda mucosa se utiliza para designar un ciclo completo de movimiento vibratorio de los pliegues vocales. Hay estudios que muestran que la velocidad de la onda mucosa aumenta con la elongación de la cuerda vocal, el incremento de flujo, aumento de la presión subglótica y con la contracción de la musculatura laríngea (Jiang, 1998; Titze et al., 1993). Al aumentar la velocidad de la onda mucosa aumenta la frecuencia fundamental. Estos conceptos son fundamentales para el diagnóstico de patologías vocales a través de técnicas estroboscópicas.

2.2.4. Mecanismo vibratorio de la cuerda vocal

A lo largo de los años, han sido varios los autores que han propuesto distintas teorías sobre cuáles son los mecanismos de producción vocal. Entre los más destacados, figura Müller (1831) que compara las cuerdas vocales con instrumentos de viento, sentando las bases para la formulación años después de la Teoría Mioelástica; Husson (1950) que expone la Teoría Neurocronáxica; Cornut y Lafón que explican en 1960 la Teoría Impusional; Van Den Berg, que en 1958 postula la Teoría Mioelástica-Aerodinámica, que será matizada años después por Perelló (1962) como Teoría Muco-Ondulatoria.

Minoru Hirano en 1974 describe por primera vez el concepto *body-cover complex* (cuerpo-cubierta) para referirse a la estructura de la cuerda vocal. Este concepto dará lugar al desarrollo de su Teoría Cuerpo-Cubierta (Hirano, 1989). Actualmente, son dos las teorías más aceptadas para describir el ciclo fonatorio (Jackson-Menaldi, 1992):

- Teoría Mioelástica-Aerodinámica. Coincide con la Teoría Muco-Ondulatoria de Perelló (1962) al suponer dos principios básicos en la producción de voz:
  - La masa de los pliegues vocales, la elasticidad cordal y la presión subglótica determinan la frecuencia fundamental.
  - La vibración vocal se explica por el principio de *Bernouilli,* que provocaría a la altura de los pliegues vocales aducidos un movimiento de aspiración hacia la línea media.

  Además, describe dos supuestos más que explicarían el acto vibratorio:

  - Las cuerdas vocales contactan en aducción en la línea media.
  - La presión subglótica del aire espirado produciría una resistencia a nivel cordal que provocaría la apertura de los pliegues, dejando salir momentáneamente el aire.

  Liberman (1968) explica de forma detallada los principios de la Teoría Mioelástica-Aerodinámica, afirmando que son dos fuerzas las que actúan en las cuerdas vocales:
  - Fuerzas aerodinámicas y aeroestáticas que desplazan los pliegues vocales desde la aducción (pre-fonación).
  - Fuerzas del tejido que devuelven los pliegues vocales a su posición de aducción.

- Teoría Cuerpo-Cubierta. Hirano (1974) describe la estructura de la cuerda vocal dividiéndola en tres planos: cubierta, capa de transición y cuerpo, atribuyendo diferentes propiedades mecánicas a cada una de ellas. Desde un punto de vista funcional, este hecho es fundamental, ya que explica cómo a través del desplazamiento de la cubierta sobre el cuerpo, se produce el movimiento vibratorio conocido como onda mucosa. Existen otros autores como Jiang et al. (2000) que diferencian únicamente dos capas en los pliegues vocales: la cubierta (formada por el epitelio y las capas superficiales y media de la lámina propia) y el cuerpo, constituido por el músculo vocal y la capa profunda de la lámina propia.

2.3. Bases físico-acústicas de la fonación

Los pliegues vocales sufren diferentes modificaciones físicas para poder adaptarse a la variabilidad de condiciones acústicas que posee la voz humana (tono, intensidad y timbre). Algunos de los factores implicados en estos procesos serían: contracción de los músculos que intervienen en el timbre e intensidad de la voz (Tiroaritenoideos), tensión o elongación de los pliegues vocales (tono), superficie de contacto de los pliegues vocales (frecuencia), presión de cierre glótico (timbre), posición de la laringe (volumen).

2.3.1. Propiedades de las ondas sonoras. El sonido

Distintas son las definiciones que en la literatura se pueden encontrar del sonido. En física, se definiría el sonido como una vibración mecánica transmitida por un medio elástico. Desde el punto de vista de las ciencias biomédicas, el sonido se definiría como la sensación producida en el órgano del oído por el movimiento vibratorio de los cuerpos, transmitido por un medio elástico, como el aire. Teniendo en cuenta los principios de la física y de las ciencias biomédicas, podría definirse el sonido como un movimiento oscilatorio, armónico o no, que se propaga por diferentes medios materiales y que se encuentra dentro de un rango de frecuencias que puede ser captado por el oído del ser humano (entre 20Hz y 20000Hz) (Frumento, 1995).

La voz es un sonido. Para entender de forma global sus características, en cuanto a su descomposición en ondas acústicas, es necesario conocer conceptos como (véase la Figura 2.5):

- Onda. Perturbación que se propaga desde el punto en que se produjo hacia el medio que rodea ese punto.
- Ciclo. Oscilación completa de una curva que se encuentra entre dos crestas o dos valles.
- Frecuencia. Número de ciclos por segundo de una onda. Se mide en hertzios (Hz). La frecuencia determinará el tono vocal (*Pitch*).
- Amplitud. Distancia máxima de una onda con respecto a su posición de equilibrio, es decir, distancia que va desde el eje de la onda a la cresta o pico de la misma. La amplitud es la que determina la intensidad (volumen) de un sonido. Amplitud e intensidad son directamente proporcionales, de modo que, a mayor amplitud de onda, mayor intensidad del sonido. La amplitud (intensidad) se mide en decibelios (dB).
- Periodo. Tiempo que transcurre entre la emisión de dos ondas consecutivas (duración de un ciclo), es decir, tiempo que tarda en pasar una onda completa por un punto de referencia.

- Timbre. Impresión psicológica del sonido que percibimos (agudos y graves). Hace referencia a los armónicos percibidos, por tanto, alude a una cualidad de la percepción de la onda, no a una cualidad física de ésta. Permite distinguir dos sonidos de igual intensidad y frecuencia emitidos por dos fuentes sonoras diferentes.

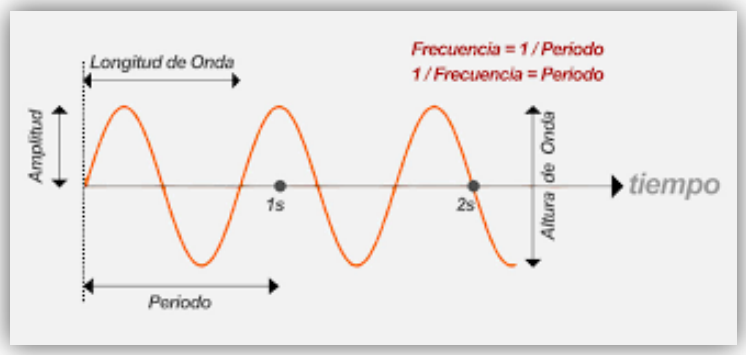

Figura 2.5. Parametrización de onda sonora (Goller, 2014).

Las oscilaciones del sonido pueden ser periódicas (sonido propiamente dicho) o aperiódicas (ruido). Dentro de las oscilaciones periódicas, si se produce un movimiento armónico diremos que el sonido es puro. Cuando el movimiento no es armónico, pero sí periódico, hablaremos de un sonido compuesto. La periodicidad de este movimiento la marcará la frecuencia fundamental, y el resto de componentes tendrán que ser múltiplos de la misma. Recibirán el nombre de armónicos, por tanto, las frecuencias que sean múltiplos de una frecuencia fundamental (Fo) definida. El primer armónico es la onda sonora simple con la frecuencia fundamental más baja, y es la frecuencia que concentra mayor cantidad de energía o potencia del sonido que se oye (Blomgren, 1998). Una onda compleja con frecuencia fundamental de 250 Hz presentará armónicos en 500 Hz, 750 Hz, 1000 Hz, etc. Es decir, los armónicos son sonidos más agudos. La voz de un varón adulto tiene una frecuencia fundamental media de aproximadamente 125 Hz, la de una mujer de 220-225 Hz. Esto hace, que percibamos más aguda (en términos de tono), la voz femenina.

2.3.2. Modelo acústico de generación de voz. Los formantes

La acústica del habla depende de múltiples factores, entre los cuales destaca el fenómeno de resonancia. El sonido producido en la laringe se transmite a lo largo del conducto vocal, donde se filtra y modifica hasta su salida por los labios. Los órganos de resonancia y de la articulación juegan, por tanto, un papel fundamental en la emisión definitiva de un sonido, de manera que crean variaciones en las frecuencias de resonancia del sonido que se conocen como formantes. Las frecuencias formantes más importantes del tracto vocal, en cuanto a la acústica de la voz hablada y cantada, son las cinco primeras (F1, F2, F3, F4 y F5).

Los formantes permiten distinguir los sonidos del habla humana sobre todo las vocales y, por tanto, son responsables del timbre de cada vocal. El formante con la frecuencia más baja o primer formante se designa F1, el segundo F2, y así sucesivamente. Los formantes dependen de la longitud total del tracto vocal y de las modificaciones morfológicas que este sufre. Al vibrar, las cuerdas vocales producen una onda compuesta que es amplificada o disminuida por los órganos de la articulación y de resonancia. La manifestación acústica en forma de onda estacionaria dentro del tracto vocal va a depender de la localización de las zonas de máxima presión de las mismas, de manera que se puede distinguir dónde se coloca la voz durante la emisión de una vocal determinada. Los dos primeros formantes (F1 y F2) permiten la identificación de las vocales. El primero (F1) controla la amplitud del sonido, y depende de la cavidad faríngea, cuanto más estrecha, mayor frecuencia, y viceversa. El segundo formante (F2) controla la inteligibilidad del sonido y está influido por la posición de la lengua, de manera que, si se eleva en la parte anterior, la frecuencia se elevará en relación directa con la altura; si es en la parte posterior, descenderá en relación inversa con la altura. El tercer formante (F3) está relacionado con la acción de los labios, de forma que, si están estirados, el valor de la frecuencia será más alto, y si están redondeados, será más bajo. Los formantes F4 y F5 varían con la anchura y longitud del tracto vocal, resultando más agudos cuanto más corto y estrecho sea (Gracida & Orduña, 2011).

Las distintas cavidades de resonancia confieren distintas cualidades a los sonidos, especialmente a las vocales. A nivel de los pliegues vocales, el sonido posee un espectro acústico regular, siendo el mismo timbre para todos los sonidos. En la parte superior de la faringe, en el velo del paladar, el espectro sonoro se expande hacia el agudo, perdiendo amplitud los sonidos graves. En la parte posterior de la cavidad bucal, el espectro se enriquece de frecuencias intermedias, y en la parte anterior, de frecuencias agudas y ruidos. Por último, fuera de la cavidad bucal, el sonido aparece con estructura de formantes (Woisard-Bassols, 2000).

2.4. Análisis acústico y estudio funcional de la voz

Thomas A. Edison anunció la invención de su primer fonógrafo, el 12 de agosto de 1877 y ese mismo año lo patentó. Este fonógrafo utilizó un sistema de grabación mecánica analógica en el que las ondas sonoras eran transformadas en vibraciones mecánicas mediante un transductor acústico-mecánico. A partir de este momento, puede considerarse que se inicia el estudio acústico de la voz en forma instrumental.

La función vocal posee diferentes particularidades que hacen necesario un estudio multidimensional que englobe diferentes parámetros para su completa valoración. Desde una perspectiva clínica, la voz humana contiene numerosos componentes acústicos que no pueden ser valorados únicamente por procedimientos perceptivos basados en la capacidad del oído humano. El estudio acústico de la voz permite establecer valoraciones directas entre rasgos acústicos medidos y percibidos. La descripción y caracterización de la patología vocal ha demandado una fuerte atención durante los últimos treinta años, tanto desde el punto de vista fisiológico y clínico (Roy et al., 2013), como desde un punto de vista funcional y biomecánico (Titze, 1994).

Este abordaje de valoración y evaluación de la voz se beneficia de los sistemas de digitalización de audio de alta calidad, y su fácil incorporación en equipos informáticos de bajo coste, disponibles en la actualidad. Esto ha permitido que surjan herramientas avanzadas para la captura y el análisis de la voz de uso clínico, de entre los que cabe destacar el sistema pionero de *Kay Elemetrics* (*Voice and Speech Lab*). El análisis acústico de la voz permite la utilización de características específicas del procesado de señal, como son los parámetros de perturbación (Jitter y Shimmer), la relación armónico-ruido HNR (*Harmonic to Noise Ratio*), la energía normalizada NNE (*Normalized Noise Energy*), o la relación excitación glótica-ruido. También permite medir alteraciones funcionales (asimetría en la vibración de los pliegues), o incluso neurológicas (hiper e hipo-tensión y temblor). Nos da la oportunidad de estudiar a partir de métodos no invasivos cualidades de la voz integradas en parámetros numéricos.

Utilizando un micrófono, un ordenador, software de grabación y algoritmos de extracción de características acústicas se puede realizar el estudio acústico de la voz. Se trata de una técnica no invasiva, que permite obtener múltiples datos de la función laríngea, articulatoria u otros aspectos vocales. Partiendo de variables cualitativas y cuantitativas se pueden obtener valoraciones más objetivas de la función vocal. Hay que tener en cuenta en todo este proceso, que no existen unos estándares de normalidad para cada una de las pruebas realizadas, ya que son múltiples las variabilidades interpersonales y también de los equipos utilizados. Al ser el sonido una onda de presión que se propaga por el aire, durante una grabación esta alcanza el micrófono, y la convierte en una señal eléctrica (onda de tensión), cuya amplitud es proporcional a la onda de presión. La memoria de un ordenador sólo contiene secuencias de números de precisión limitada, por lo que la señal sonora (continua) no puede almacenarse directamente en él, sino que tiene que someterse a procesos de muestreo y cuantificación. La señal capturada por un micrófono (analógica) para poder ser procesada y analizada necesita ser digitalizada, de manera que sus valores continuos puedan transformarse en una serie de muestras o números discretos. Esto se consigue a través de los procesos de muestreo y cuantificación (Cobeta & Núñez, 2013).

El teorema de Nyquist demuestra que la reconstrucción exacta de una señal periódica (como el sonido) a partir de sus muestras, es matemáticamente posible si la señal está limitada en banda y la tasa de muestreo es superior al doble de su ancho de banda, es decir, que no es necesaria una frecuencia de muestreo infinita para poder obtener toda la información de una señal.

Las variaciones y perturbaciones en la intensidad y la frecuencia, así como la presencia de ruido espectral son aspectos que están presentes en la mayor parte de las voces disfónicas. Se pueden clasificar, por tanto, parámetros acústicos en función de la información que aporten tras su estudio, encontrando:

- Parámetros característicos del propio comportamiento vocal, presentes en situaciones de excitación vocal: frecuencia o tono, amplitud o intensidad y espectro (relación entre la intensidad y la frecuencia).
- Parámetros presentes en el estudio del tracto vocal: ruido espectral, frecuencia de los formantes y envolvente del espectro.

El proceso de grabación, extracción y selección de características y clasificación de señales de voz constituye un sistema CAD (*Computer-Aided Diagnosis*). Estas técnicas buscan maximizar la información que se puede extraer automáticamente de pruebas médicas (imágenes, grabaciones de voz, vídeos…) mediante cálculos con modelos cuantitativos (Calle-Alonso et al., 2013).

2.4.1. Parámetros de frecuencia

La frecuencia fundamental (Fo) representa el número de veces que las cuerdas vocales se abren y cierran por segundo, y se expresa en ciclos por segundo o Hertzios (Hz) (véase la Figura 2.6). La Fo es el nivel óptimo de vibración en el cual la voz produce una frecuencia confortable sin la menor tensión laríngea. Corresponde con el armónico más bajo en la señal acústica. Descrita desde un punto de vista subjetivo, la Fo se puede calificar como aguda, grave o inestable (Heuillet-Martin et al., 2007).

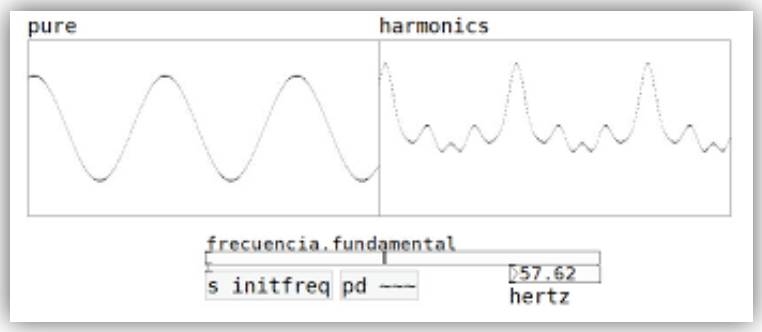

Figura 2.6. Frecuencia fundamental y harmónicos (Goller, 2014).

Los valores rango definidos para la Fo dependen de la edad y el sexo, y están relacionados con la longitud de los pliegues vocales. Según Farías (2016) y Jackson-Menaldi (2002) en un recién nacido, la Fo alcanza un valor en torno a los 500 Hz (pliegues vocales 2 milímetros), hasta los 8 años 300 Hz (pliegues vocales 6-8 milímetros) en una mujer adulta 220-225 Hz (pliegues vocales 10-18 milímetros) y en un hombre adulto 125 Hz (pliegues vocales 16-25 milímetros). La duración de cada ciclo que compone la Fo se denomina periodo (T), y se mide en segundos. La frecuencia y el periodo son inversamente proporcionales: T = 1/Fo, Fo = 1/T. La Fo está influida por la estructura y longitud anatómica de los pliegues vocales, de forma que, al aumentar la masa de éstos, la frecuencia disminuye, adquiriendo la voz un tono más grave. Sin embargo, los aumentos de tensión y longitud de las cuerdas vocales, y los de la presión subglótica, producen aumentos en la frecuencia de vibración, dando lugar a tonos más agudos. Estas diferencias en los valores de la frecuencia fundamental se pueden observar en parámetros relacionados con la misma: Fhi (frecuencia fundamental más alta), Flo (frecuencia fundamental más baja), PFR (rango fonatorio en semitonos, rango entre Fhi y Flo).

La perturbación de frecuencia o Jitter (Orlikoff, 1990) es el promedio de variación entre periodos, medido en microsegundos ciclo a ciclo. Es uno de los más empleados en clínica, y representa una medida de la estabilidad de la fonación. Es por esto, por lo que para su medición hay que desechar el principio y el final de señal analizada, ya que son las zonas de mayor inestabilidad. La entonación de una frase puede hacer que la Fo varíe, pero el Jitter sólo recogería aquellas variaciones que se producen entre un ciclo vocal y el siguiente.

Pueden diferenciarse varios tipos de medidas de la perturbación de frecuencia:

- Jitter absoluto (*Absolute Jitter)*, que mide la variación entre ciclos de la frecuencia fundamental de la muestra analizada, y se expresa en unidades de tiempo (microsegundos(μs)). Al ser la Fo diferente entre mujeres y hombres, este parámetro también se verá afectado por la variable sexo. Expresado como:

$$Jitter(absoluto) = \frac{1}{N-1}\sum_{i=1}^{N-1}|T_i - T_{i+1}|,$$

  donde T_i es la duración de cada ciclo de la Fo, y N es el número de ciclos.

- Jitter relativo (*Jitter Percent*), que mide la variabilidad relativa ciclo a ciclo. Se obtiene dividiendo el Jitter medio absoluto entre el tono medio. Se expresa en porcentaje (ratio), y no se ve afectado por la Fo, y no existen diferencias de género. Se define como:

$$Jitter(relativo) = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}|T_i - T_{i+1}|}{\frac{1}{N}\sum_{i=1}^{N-1}|T_i|}$$

Existen otras medidas de perturbación de frecuencia a menudo utilizadas en la literatura por ser menos sensibles a posibles alteraciones erráticas entre dos ciclos consecutivos (variaciones de tono):

- RAP (*Relative Average Perturbation*), que promedia la diferencia de tres periodos consecutivos:

$$RAP = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}\left|T_i - \frac{1}{3}\sum_{n=i-1}^{i+1}T_n\right|}{\frac{1}{N}\sum_{i=1}^{N}|T_i|}\cdot 100$$

- PPQ (*Pitch Period Perturbation Quotient*), que mide la variabilidad de la frecuencia fundamental en grupos de cinco ciclos:

$$PPQ = \frac{\frac{1}{N-1}\sum_{i=2}^{N-2}\left|T_i - \frac{1}{5}\sum_{n=i-2}^{i+2} T_n\right|}{\frac{1}{N}\sum_{i=1}^{N}|T_i|} \cdot 100$$

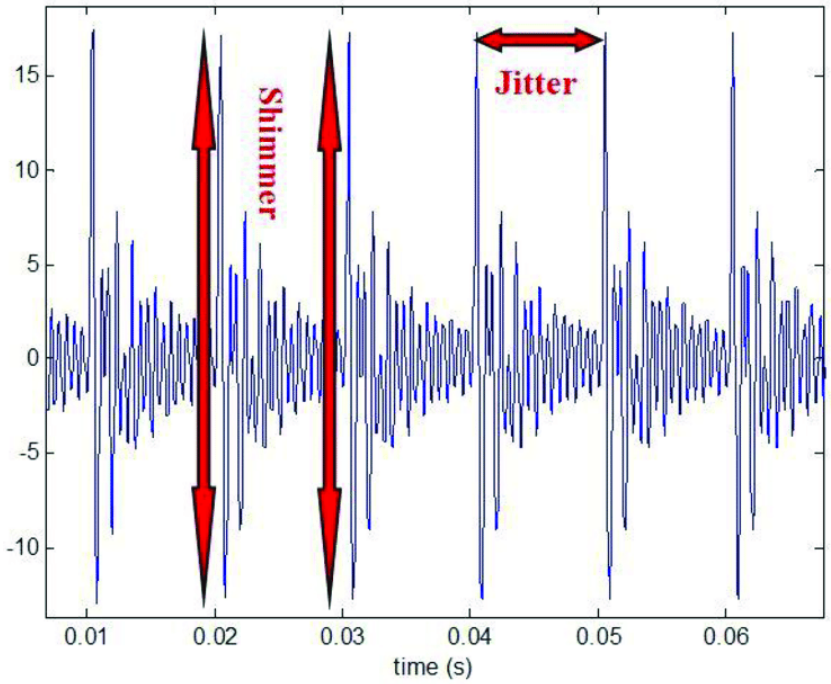

Figura 2.7. Jitter y Shimmer en una señal acústica (Teixeira & Gonçalves, 2014).

2.4.2. Parámetros de perturbación de amplitud

Se define la intensidad de un sonido como la amplitud de la variación de la presión sonora producida al transmitirse en el medio aéreo, y se expresa en decibelios (dB). Se traduce en un registro oscilográfico por la amplitud de las ondas del trazado.

El volumen es la sensación psico-acústica del fenómeno físico de la intensidad. Su valor va a depender de la amplitud de la vibración de los pliegues vocales y de la presión subglótica, de manera que cuando éstos aumentan, también lo hace la intensidad de la voz. La intensidad vocal según Husson (1950) será de 10 centímetros de agua en una conversación tranquila (30dB), de 60 cm de agua en cantos de salón (60dB), de 100 cm de agua en voces de advertencia (70dB), de 160 cm de agua en los picos de intensidad de un orador en una exposición (80dB) y de 360 cm de agua en un primer tenor al máximo de su potencia (120dB).

Factores como un cierre glótico incompleto, pliegues vocales rígidos o soporte respiratorio inadecuado pueden afectar de forma negativa a la intensidad de la voz.

Por otra parte, se puede dar un aumento de intensidad en situaciones en las que aumenta la vibración de las cuerdas vocales y la presión subglótica, y una disminución en el caso de soplos fonatorios deficientes (soporte respiratorio deficiente), cierre glótico incompleto o cuerdas rígidas (onda mucosa disminuida).

La perturbación de amplitud o Shimmer mide la variabilidad de la amplitud ciclo a ciclo. Representa una medida de la estabilidad de la fonación (Baken & Orlikoff, 2000) (véase la Figura 2.7). Tiene una relación inversa con la intensidad vocal, de manera que disminuye al aumentar la intensidad. Al igual que el Jitter puede sufrir variaciones relacionadas con alteraciones en los pliegues vocales.

Existen distintos parámetros de medida Shimmer:

- Shimmer local (*Local Shimmer*), que mide en dB la variabilidad de la amplitud (pico a pico) de la onda de voz analizada. No presenta diferencia entre sexos, pero sí entre distintos sujetos, lo que lo convierte en un parámetro muy variable con un rango de normalidad muy amplio. Se define:

$$Shimmer\ (dB) = \frac{1}{N-1}\sum_{i=1}^{N-1}\left|20\log\left({A_{i+1}}/{A_i}\right)\right|,$$

  donde A_i es la amplitud pico a pico, y N es el número de periodos de la frecuencia fundamental.

- Shimmer relativo (*Shimmer Percent*), que determina la variabilidad entre ciclos consecutivos de la amplitud. No presenta variaciones respecto al género. Se obtiene dividiendo el Shimmer medio absoluto entre la amplitud media, y se expresa en porcentaje:

$$Shimmer\ (relativo) = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}\ |A_i - A_{i+1}|}{\frac{1}{N}\sum_{i=1}^{N} A_i}$$

- APQ (*Amplitude Perturbation Quotient*), que analiza la variabilidad relativa de la amplitud de la onda analizada en grupos de tres en tres, cinco en cinco u once en once periodos. La fórmula siguiente representa grupos de tres en tres:

$$APQ = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}\left|A_i - \frac{1}{3}\sum_{n=i-1}^{i+1} A_n\right|}{\frac{1}{N}\sum_{i=1}^{N} A_i} * 100$$

- sAPQ (*Smoothed Amplitude Perturbation Quotient*), que evalúa la variabilidad relativa de la amplitud en ciclos de 55 en 55.

$$sAPQ = \frac{\frac{1}{N-1}\sum_{i=2}^{N-2}\left|A_i - \frac{1}{55}\sum_{n=i-2}^{i+2} A_n\right|}{\frac{1}{N}\sum_{i=2}^{N} A_i} \cdot 100$$

2.4.3. Parámetros de ruido

El ruido puede tener diversos orígenes en el análisis de una señal de voz: variaciones en la frecuencia o amplitud, turbulencia, subarmónicos o rupturas de voz. La razón armónicos/ruido determina el cociente entre el componente periódico (armónicos) y la energía aperiódica (ruido) y se expresa en porcentaje. Las propiedades de la señal acústica vocal pueden contribuir a la aparición de ruido en el espectro (perturbaciones de la fuente, cambios en la Fo, transiciones en los formantes, etc.), por lo que no se puede definir de forma concluyente la presencia o no de ruido en la señal a partir de la observación del nivel de ruido del espectro. Existen distintas fuentes de ruido espectral según el autor que realice el estudio, por lo que los resultados obtenidos sólo atienden al criterio establecido por el investigador previamente.

La relación ruido/armónico se determina a partir del parámetro *Noise to Harmonic* Ratio (NHR). Se expresa en tanto por ciento y permite realizar una evaluación general del ruido en una señal analizada. Por otra parte, el *Signal to Noise Ratio* (SNR) muestra la relación entre la energía de la señal y la energía de ruido. A pesar de ser el parámetro menos utilizado en la literatura, destaca su utilidad en la detección del tipo y la severidad de la disfonía (Yumoto et al., 1982). Finalmente, el *Normalice Noise Energy* (NNE) transforma los valores de la intensidad de ruido encontrados en una muestra analizada en valores de una distribución normal. Presenta valores negativos, siendo los más cercanos a cero los que se identifican con rangos patológicos (Cobeta & Núñez, 2013).

2.4.4. Parámetros acústicos no lineales

La señal acústica vocal se caracteriza por ruido y no periodicidad (Little et al., 2007). Esto ha hecho que recientemente se hayan desarrollado distintos algoritmos para el análisis acústico de la voz que están basados en métodos no lineales de procesamiento (Duffy 2005; Little, 2007). Estos parámetros han sido eficientes en la evaluación de rasgos de disfonía donde ha existido una alteración en alguna de las estructuras u órganos que forman el aparato fonador. Las características acústicas de una voz disfónica han podido ser estudiadas a partir de algoritmos que miden dimensiones de correlación y sucesiones pseudoperiódicas (Zhang & Small, 2006). La utilización de parámetros no lineales como *Detrended Fluctuation Analysis* (DFA), que mide las fluctuaciones sin tendencia, *Recurrence Period Density Entropy* (RPDE), que mide las desviaciones con respecto a un comportamiento periódico perfecto, o *Pitch Period Entropy* (PPE), que mide la entropía del tono, han resultado muy útiles para detectar alteraciones de la voz (Little et al., 2007).

2.4.5. Estudio funcional de la voz

En la valoración funcional de la voz se presta atención a distintas variables subjetivas y objetivas que engloban todas las dimensiones de la voz de un sujeto, incluyendo la repercusión en su vida profesional, social y emocional. Deberá incluir un estudio orgánico y funcional, así como del comportamiento fonatorio. La normalidad vocal no existe, las referencias a una voz «buena» son esencialmente culturales y se relacionan con las expectativas sociales y personales de cada uno (Woisard-Bassols, 2000). La emisión vocal puede ser valorada a partir de:

- Ataque vocal, que valora la forma de inicio del sonido al articular una palabra, y está relacionado con la configuración glótica en el momento de la emisión. Puede ser normal, duro (inicia la fonación con una compresión medial intensa, típico de disfonías hipertónicas) o soplado (se percibe el aire antes de la emisión del tono estable, antes del cierre glótico completo).
- Estabilidad de la emisión vocal. Relacionada con aspectos emocionales, de entrenamiento vocal e incluso déficits neurológicos. Una sustentación no adecuada puede suponer la alteración de algunos de estos parámetros.
- Registro vocal. Hace referencia a los distintos tonos que puede emitir el individuo con comodidad. Los principales registros son: basal, modal y alto.
- Prosodia, que valora la información no verbal que puede dar el discurso de un individuo. Está relacionada con las variaciones de la frecuencia en la voz hablada, y en ella podemos distinguir voces monótonas, inestables o temblorosas.

Por otra parte, el hábito fonatorio se puede analizar con diferentes pruebas. Las más importantes son:

- Espirometría. Valora si existe o no insuficiencia en la mecánica respiratoria. Las variables más importantes que recoge son (García Ríos et al., 2013):
  - Capacidad vital (CV). Volumen de aire expulsado en una espiración forzada tras una inspiración forzada. Se mide en litros, y debe ser similar al FVC.
  - FEV1. Volumen máximo de aire espirado en el primer segundo.
  - Capacidad vital forzada (FVC). Volumen total de aire espirado desde la inspiración máxima a la espiración máxima.
  - Índice de Tiffeneau (FEV1/FVC). Porcentaje del volumen total espirado en el primer segundo. Sus valores normales están entre el 70-75%, y aporta información del patrón respiratorio que presenta una persona.
- Hábito respiratorio. Variable cualitativa que nos informa sobre el patrón respiratorio utilizado durante la mayor parte de la emisión vocal. Se recogen tres opciones: abdominal, costal media o torácica y costal alta o clavicular.

- Coordinación fono-respiratoria. La relación inspiración/espiración del acto ventilatorio en el recambio gaseoso es de 1:2. En una conversación 1:6, y en el canto puede llegar a ser 1:15. Con relación a la velocidad del soplo, presión subglótica, músculos espiratorios y la capacidad de cierre glótico se obtendrán distintos tipos de voces. Por tanto, la coordinación fono-respiratoria se convierte en el resultado de la integración de las fuerzas espiratorias, miolásticas de la laringe y musculares de la articulación y la respiración. A partir de su valoración se comparan la eficacia del cierre glótico y el sistema respiratorio. Los parámetros vocales más útiles para dicha valoración serían:

  - Tiempo máximo fonatorio (TMF). Mide el tiempo máximo de emisión de una vocal (generalmente la /a/ o la /e/) en una frecuencia e intensidad cómoda, tras una inspiración profunda. Se utiliza también para la valoración del tiempo de emisión del fonema /z/. Su valor varía en función de la capacidad pulmonar, la intensidad del sonido y la eficiencia glótica. Se repite tres veces, y se toma como valor de referencia el mejor tiempo realizado. Los valores normales según Farías (2010) en hombres son de 25 a 35 segundos, mientras que en las mujeres el valor normal se establece entre 15–25 segundos. El tiempo aproximado en los niños es de 10 segundos (Colton et al., 2011). Existen distintas fórmulas que posibilitan la estimación del TMF en hombres y mujeres. En el caso de los primeros, el TMF = 0,67 × CV / 110, y en el caso del sexo femenino el TMF = 0,59 × CV / 110, donde CV es la capacidad vital medida por espirometría (Jackson-Menaldi & Núñez, 2013). Se considera un valor patológico por debajo de 10-15 segundos, y se observa en las disfonías con un defecto de cierre glótico, como son las parálisis laríngeas en abducción, los sulcus y los nódulos y pólipos vocales (Shigemori, 1977). Se puede dar el caso de TMF alargado en disfonías hipertónicas (Freeman, 2000).

  - Tiempo máximo espiratorio (TME) o tiempo máximo de soplo (TMS). No valora directamente un aspecto relacionado con la fonación, pero a partir de él se calculará un parámetro de relación entre la fonación y la espiración (cociente fonorrespiratorio). Se evalúa a partir de la emisión de un soplo espiratorio durante el máximo tiempo posible con la emisión de un fonema /s/. Se repite tres veces, y se toma como valor de referencia el mejor tiempo realizado.

  - Índice s/z (o índice s/e) o Índice de Fonación (IF). Se conoce también como cociente fonorrespiratorio clínico, y se interpreta como un indicador de la eficiencia de cierre glótico. Se define cociente s/z como el valor (medido en segundos) que resulta de dividir el TME (o TMS) para la /s/ entre el TMF para la /z/. El TME de la /s/ depende exclusivamente de la función pulmonar, por tanto, este índice relaciona la función pulmonar y la laríngea. Los valores normales se sitúan entre 1,4-1,5, rangos superiores se relacionan con incompetencia glótica por defecto de cierre, o lesiones en el borde libre de los pliegues vocales.

Los cocientes cercanos a la unidad indican problemas respiratorios, con una capacidad vital que puede estar reducida, y una deficiente fuerza espiratoria.

- Cociente fonatorio (CF). Es la medida del control respiratorio y de la eficiencia del cierre glótico. Su valor es el resultado de dividir la CV entre el TMF. Los valores estándar son de 145 ml/s para los hombres y 137 ml/s para las mujeres. Un CF inferior a 69 ml/s o por encima de 307 ml/s para los hombres, y menor de 78 ml/s o mayor de 241 ml/s para las mujeres, debe ser considerado anormal. El enfisema pulmonar, asma, obstrucciones nasales, neoplasias laríngeas, sinequias laríngeas, paresias o parálisis de los pliegues vocales son patologías relacionadas con CF anormales (Jackson-Menaldi & Núñez, 2013).

- Alineación corporal. La estática corporal y el equilibrio muscular favorecen la producción de voz con el menor esfuerzo posible. La postura de la columna cervical, sin desviaciones, posibilita el movimiento de la laringe y la correcta proyección de la voz. Se prestará especial atención por tanto a la correcta verticalidad del cuerpo.

La valoración subjetiva de la voz por parte de la persona que padece una patología vocal forma parte de la evaluación clínica en la práctica profesional. La cuantificación de la calidad de vida de las personas que conviven con voces patológicas y el impacto que la enfermedad tiene en su desarrollo diario, ha impulsado el desarrollo de numerosos cuestionarios de medida. Se pueden identificar hasta doce cuestionarios de percepción de disfonía (Zraick et al., 2008), pero sólo el que mide el Índice de Incapacidad Vocal o *Voice Handicap Index* (VHI) ha sido traducido, validado y adaptado a la lengua castellana (Núñez-Batalla et al., 2007).

El cuestionario VHI fue desarrollado por Jacobson et al. (1999), a partir de la valoración de 30 ítems organizados en tres subescalas: física, emocional y funcional. El cuestionario recoge y cuantifica el impacto que el trastorno vocal tiene en la persona afectada en las tres esferas evaluadas. Cada pregunta se responde con una puntuación de 0 a 4, siendo 0= nunca, 1= casi nunca, 2= algunas veces, 3= casi siempre y 4= siempre. El valor máximo del VHI es de 120 puntos. Existe una versión corta de sólo 10 ítems, en los que se siguen valorando las tres esferas, cuya fiabilidad y potencia ha sido demostrada (Rosen et al., 2004).

La Valoración Perceptual de la voz (valoración psicoacústica) está incluida en la mayor parte de protocolos de evaluación de clínica de personas afectadas por trastornos vocales (Oates, 2009; Rodriguez-Parra et al., 2009; Webb et al., 2004). Hirano (1981) definió una escala de valoración perceptiva de la voz, con el fin de unificar criterios de evaluación perceptuales de voz, el sistema GRABS (o GRBAS). El acrónimo GRABS se corresponde en inglés con las palabras: *Grade* (grado), *Rough* (rasposidad, ronquera), *Astenic* (asténica, cansada), *Breathy* (aérea, soplada), *Strain* (tensa, constreñida).

Cada uno de estos atributos puede tener un valor de 0 a 3, donde 0 es normal, 1 ligero, 2 moderado y 3 extremo. De los parámetros analizados por la escala GRABS, cabe señalar que las alteraciones descritas por las letras A y S se asocian generalmente a lesiones funcionales, y las recogidas por la R y la B lo hacen a lesiones orgánicas. La G indica el grado general de afectación de la voz, donde el 0 corresponde a una voz normal. La R valora el grado de disfonía, describe si la voz es rasposa o ronca. La A indica el grado de debilidad o fatiga que se aprecia en la emisión vocal, reflejando la incapacidad de fonación durante periodos prolongados de tiempo sin necesidad de cambiar el timbre vocal. La voz se percibe más grave, ya que suele existir dolor acompañado de un cambio en el timbre. Existen ocasiones, en las que se trata de forzar una voz más grave, disminuyendo la Fo. Estas actuaciones sostenidas de descenso y ensanche de faringe, terminan produciendo alteraciones en las estructuras vocales, y agotamiento del acto fonatorio. Este hecho fue descrito y denominado por Koufman (Koufman, 1988) como "Síndrome de *Bogart-Bacall*". La B hace alusión a la voz soplada. Indica un defecto en el cierre glótico. Por último, la S hace referencia a una voz tensa, o constreñida. Corresponde a la hiperfonación o tensión excesiva de la laringe.

Capítulo 3. Programa de prevención de riesgos en la voz

El diccionario de la Real Academia de la Lengua (RAE) en su actualización de 2018, define voz en sus dos primeras acepciones como: “1. f. Sonido producido por la vibración de las cuerdas vocales. 2. f. Calidad, timbre o intensidad de la voz” (Diccionario RAE, 2018). La primera de ellas menciona la voz como cualidad humana, donde la fuente primaria de producción del sonido definido como voz son las cuerdas vocales. Sin embargo, en el mecanismo de producción de voz se ven implicadas además otras estructuras corporales como son los pulmones, el diafragma, la laringe, las cavidades resonadoras y las articuladoras. La segunda acepción del diccionario de la RAE alude a tres propiedades del sonido que caracterizan perceptualmente a la voz, y que permiten distinguir entre sonidos provenientes de fuentes diferentes, así como su agudeza o gravedad.

El interés por el estudio de la voz humana y sus trastornos ha crecido en las últimas décadas. La atención de los problemas vocales ha pasado de ser un lujo al alcance de unos pocos, a incluirse dentro de un concepto de salud integral del individuo (Casado et al., 2001).

En el presente capítulo se abordarán aspectos relacionados con las características de la voz normal y patológica, las características vibratorias de nódulos, pólipos y edema de Reinke, el uso profesional de la voz, la salud vocal del docente, los niveles de prevención y el programa que se propone sobre prevención de disfonías en docentes.

3.1. La voz normal y la voz patológica

La cualidad humana de producción de voz voluntaria como instrumento de comunicación y lenguaje, supone un importante y amplio campo de estudio. El concepto de voz normal, definido a partir de criterios objetivos y absolutos no existe como tal. Son diferentes los condicionantes que pueden hacer que una voz suene de una forma u otra, haciendo ilimitada la variedad de voces que se pueden encontrar. Los estándares que cumple una voz sana o normal son amplios, de manera que, determinar una definición de voz normal resulta complicado (Moore, 1971). Las referencias a una voz normal vienen determinadas por aspectos culturales, relacionados con expectativas sociales y personales (Avilés et al., 2014). Autores como Casado et al. (2001) señalan que el concepto que más se aproxima al de voz normal, sería el que describe la voz de un paciente no fumador, sin historia previa de patología vocal, sin problemas respiratorios, neurológicos, nasosinusales o faringolaríngeos previos, sin antecedentes de alergia, con una audición normal y sin uso profesional de la voz.

Cobeta et al. (2013c) utilizando como base los estudios de Johnson et al. (1965) establecieron una serie de criterios generales a partir de los que se podría hablar de voz normal. Estos serían:

- Timbre agradable, donde debiera existir cierta sonoridad musical y ausencia de ruido o atonalidad.
- Tono adecuado a la edad y el sexo de la persona emisora de voz.
- Volumen adecuado, de manera que pudiera escucharse en situaciones de habla ordinaria, y no resultara desagradable por su elevada intensidad.
- Flexibilidad adecuada, en la que no existieran grandes variaciones en el tono y el volumen durante la enfatización o expresión de sentimientos, ni en los rasgos suprasegmentales del habla: ritmo, prosodia y entonación (Dodero et al., 2005).

La diferenciación entre distintos tipos de voces, existiendo voces infantiles, voces de niño, de niña, voces de hombre y de mujer, voces de anciano fue tratada por Moore (1971). Además, entre estos grupos, podían a su vez, coexistir voces normales y anormales. La clasificación de la voz en un tipo u otro será realizada por la persona que realice su valoración. Este aspecto estará determinado por cada observador en base a sus criterios culturales, educativos, ambientales, de conocimiento vocal y factores similares (Cobeta, et al., 2013c).

Se considera voz patológica aquella en la que se evidencia un trastorno en su timbre, tono, intensidad o flexibilidad, en relación con la de otras personas de similares características de sexo, edad y grupo cultural (Aronson, 1990). Los trastornos de la voz han pasado de ser considerados una consecuencia inevitable de una grave enfermedad subyacente o un problema menor (Stemple, 1999), a ser percibidos por el individuo como la alteración de un parámetro vital que le causa malestar o incapacidad de adaptación a su medio social y/o laboral (Casado et al., 2001).

El diccionario RAE (2018) define “disfonía” en su primera acepción como: “1. f. Med. Trastorno cualitativo o cuantitativo de la fonación por causas orgánicas o funcionales”. Si tenemos en cuenta el carácter cualitativo de la definición, podremos decir, que el término disfonía hace referencia a una alteración de las cualidades de la voz. Al incorporar como causa del trastorno aspectos orgánicos y funcionales, quedarían encuadrados en la definición aquellos aspectos referentes a la función y participación que diversos órganos y estructuras tienen durante la fonación.

La etiología de los procesos patológicos que afectan a las cuerdas vocales es variada, por lo que su estudio y clasificación no ha resultado. No existe una nomenclatura estandarizada a la hora de clasificar las patologías vocales, existiendo en la literatura médica diversas propuestas. Una clasificación ampliamente utilizada es la propuesta por Cobeta et al. (2013c), que divide los trastornos de la voz en: funcionales, orgánicos y orgánico-funcionales.

3.1.1. Disfonías Funcionales

Se puede considerar disfonía funcional aquella en la que, al observar y explorar la laringe, no se evidencia razón orgánica (morfológica) que la justifique (Koufman & Blalock, 1991). Gran parte de las disfonías funcionales son consecuencia de un abuso o mal uso de la voz, en el que el incremento de la tensión muscular durante la fonación da lugar a una serie de patrones estructurales que pueden desembocar en una patología orgánica.

Se puede hablar de disfonías hipercinéticas en aquellos casos en los que se produce una contracción excesiva de la musculatura intrínseca de la laringe durante la fonación (Cobeta et al., 2008). Este tipo de disfonías viene acompañado de una clínica que el paciente suele definir con carraspeo frecuente, sensación de cuerpo de extraño, mejoría con el reposo vocal, y aparición y empeoramiento tras el esfuerzo vocal. Desde el punto de vista del observador, se puede apreciar respiración superior y/o mala coordinación fono-respiratoria. Suele verse especialmente en mujeres, niños o docentes, y aparece asociada a una disminución de la intensidad, del tono y del tiempo máximo fonatorio, con un timbre incómodo, que puede ser nasalizado (Avilés et al., 2014).
Existen dos formas de disfonía por tensión muscular: primaria y secundaria. En la primaria, se pueden observar durante la fonación movimientos laríngeos excesivos, atípicos o anómalos, sin que haya una causa obvia neurológica ni psicógena. La secundaria aparece como consecuencia de una lesión orgánica (Cobeta et al., 2013b). Según el grado de tensión muscular glótica y supraglótica, los patrones de comportamiento muscular que pueden verse en la laringe son:

- Contracción isométrica. Caracterizada por una hipertonía de los músculos intrínsecos y los suprahiodeos. Se produce un defecto del cierre glótico posterior.
- Contracción medial glótica o supraglótica. Se puede observar hipertonía de las bandas ventriculares, con aproximación de las mismas a la línea media.
- Contracción anteroposterior del plano gótico. Se produce una disminución del espacio entre la epiglotis y los Aritenoides, con acortamiento notable del diámetro anteroposterior laríngeo durante el habla. Se producen tonos vocales más graves, y existe dificultad para la emisión de sonidos agudos.
- Contracción esfinteriana. Se caracteriza por la aproximación de los cartílagos Aritenoides hacia la cara laríngea de la epiglotis.

Diversas causas pueden dar lugar a este tipo de disfonías, pudiendo agruparse en tres categorías (Cobeta et al., 2013b). La primera engloba aquellos aspectos relacionados con factores psicológicos o de personalidad, como introversión, neuroticismo (social), ansiedad, estrés reactivo o depresión (Roy & Bless, 2000). La segunda es aquella que en la que la disfonía se produce por el abuso y mal uso vocal. Está influida por la utilización inadecuada de los músculos voluntarios de la fonación, conducta que fomenta la adquisición de técnicas vocales incorrectas.

En la tercera categoría se consideran como causa de disfonías hipertónicas aquellos trastornos orgánicos, que, por su naturaleza, indujeran conductas de compensación vocal como lesiones vocales orgánicas, reflujo faringolaríngeo, trastornos hormonales, envejecimiento o infecciones respiratorias de vías altas (Belafsky et al., 2002).

La hipofunción laríngea o disfonía hipofuncional es un trastorno vocal caracterizado por un defecto del cierre glótico, que se produce como consecuencia de una disminución de la tensión muscular laríngea. La fonastenia o fatiga vocal, es la causa más frecuente de este problema, donde la voz aérea o soplada es protagonista de la fonación. La atrofia y/o la parálisis de cuerdas también pueden dar lugar a este trastorno. Los pacientes con este padecimiento refieren aumento de secreciones, con tendencia al carraspeo, y tono más grave de la voz.

La presbifonía se encuadra dentro de las hipofonías por presentar características similares en cuanto a la calidad de la voz. Se conoce como presbifonía al deterioro vocal que aparece en una persona como consecuencia del proceso de envejecimiento que experimenta el cuerpo humano a partir de la tercera edad de la vida. Los cambios producidos por la edad en diferentes estructuras y sistemas corporales están implicados en la producción vocal, contribuyendo en numerosas ocasiones a la aparición de perturbaciones vocales. Entre los más importantes destacan los producidos en el sistema respiratorio, muscular, óseo, neurológico, auditivo u hormonal.

Por último, la puberfonía o trastorno de la mutación, y la disfonía tonal pueden considerarse disfonías funcionales, ya que su padecimiento no implica la presencia de patología orgánica, sino esfuerzos musculares que pueden desembocar en daño estructural.

3.1.2. Disfonías Orgánicas

Las disfonías orgánicas son aquellos trastornos de la voz motivados por una patología de característica estructural, generalmente, en alguno de los órganos de la fonación. Atendiendo a su presentación, pueden dividirse en congénitas y adquiridas.

Las disfonías orgánicas congénitas englobarían a aquellos trastornos vocales que están presentes en el momento del nacimiento. Entre ellos destacan la laringomalacia, la sinequia congénita, el sulcus (bolsillo, estría (vergeture)), el puente mucoso y el quiste subepitelial.

La laringomalacia se produce al colapsarse los tejidos de la laringe durante la inspiración. Al disminuir su diámetro, se produce un sonido silbante llamado estridor, que aparece desde los primeros días de vida, y empeora con la alimentación, el llanto y al tumbar al recién nacido. Su evolución suele ser hacia la mejoría de forma espontánea a lo largo del primer año de vida.

En los recién nacidos que la presentan se han descrito algunas anomalías de la forma de diversas estructuras de la laringe, tales como la epiglotis, los Aritenoides o los ligamentos Aritenoepiglóticos.

La sinequia vocal es la unión fibrosa de las dos cuerdas vocales. Suele presentarse más frecuentemente en la parte anterior de los pliegues vocales (Cuestas et al., 2013). La sinequia congénita es una entidad que puede tardar en diagnosticarse, debido normalmente a su reducido tamaño y a la escasa sintomatología que la suele acompañar. Debido a la afectación que esta produce en la biomecánica vocal, es probable su presencia en pacientes que presentan nódulos de repetición.

El *Sulcus Vocalis* es una lesión que se manifiesta paralela al borde libre de los pliegues vocales. Se caracteriza por la presencia de tejido cicatricial en la mucosa superficial del espacio de Reinke (Farías, 2014; Ford, 1999). Según haya afectación o no de la onda mucosa, se pueden distinguir varios tipos de lesiones. La primera de ellas, no patológica, presenta una lámina propia normal, y se denomina sulcus fisiológico (tipo 1). La segunda, patológica, se caracteriza por la ausencia de la capa superficial de la lámina propia, con una depresión lineal o estría (vergeture) (tipo 2), o una cavidad que puede extenderse hacia el ligamento vocal o el músculo Tiroaritenoideo (tipo 3). Estas alteraciones de la lámina propia en los distintos tipos de lesiones suponen una disminución de la onda mucosa debido a la rigidez del tejido lo que provoca finalmente una acentuada disfonía (Ford et al., 1996).

Los puentes mucosos vocales son lesiones situadas de forma paralela al borde libre de la cuerda vocal, compuestas por epitelio pluriestratificado que, a modo de brida, se sitúan entre dos puntos distantes entre sí de la cuerda vocal, sin fijación a lo largo de su recorrido. Suelen acompañarse de otras lesiones como quistes, estrías o hendiduras (Díaz de Cerio et al., 2010).

Los quistes epidérmicos vocales son lesiones encapsuladas revestidas de epitelio, situadas en la lámina propia de la cuerda vocal. Se observan como un aumento de volumen fusiforme de color amarillo o blanco nacarado dentro de la lámina propia (Chuang et al., 2017)

Las disfonías orgánicas adquiridas incluirían a aquellas lesiones traumáticas, inflamatorias, neoplásicas, endocrinas, neurológicas y psicógenas.

3.1.3. Disfonías orgánico-funcionales

Las disfonías orgánico-funcionales comprenden una serie de entidades clínicas en las que se observan lesiones benignas mínimas en una o las dos cuerdas vocales. El carácter multifactorial de la etiología de la patología vocal conlleva no poder asociar la aparición de lesiones vocales a un aspecto determinado. Sin embargo, la persistencia de un patrón de fonación anómalo, en el que está presente el mal uso y/o abuso de la voz, está relacionado con la aparición de estas patologías en la mayoría de los casos. Esto resulta relevante a la hora del enfoque terapéutico.

Las lesiones exudativas del espacio de Reinke son afecciones que tienen como características comunes el carácter benigno y la localización. Se sitúan en la capa superficial de la lámina propia y se extienden a lo largo de la cuerda vocal. El concepto de lesión exudativa del espacio de Reinke surge a raíz de la publicación de diversos estudios, en los que se afirma y comprueba que ciertas lesiones benignas de las cuerdas vocales, como son nódulos, pólipos, pseudoquistes serosos y edema de Reinke, comparten características comunes en cuanto a su fisiopatología. Este hecho, hace que su diferenciación anatomopatológica en ocasiones sea complicada, dificultando así el diagnóstico definitivo (Dikkers & Nikkels, 1995; Marcotullio et al., 2002).

Los nódulos vocales son lesiones benignas que se caracterizan por una inflamación en el espacio de transición entre el tercio medio y anterior cordal (Altman, 2007). Suelen aparecer como lesiones bilaterales simétricas en la zona de mayor impacto vocal, adoptando entonces el nombre de nódulos *kissing*. Su fisiopatología está relacionada con microtraumatismos repetidos en la mucosa de la cuerda vocal por esfuerzos vocales (véase la Figura 3.1). Esto produciría congestión, hemorragia y edema, que posteriormente derivaría en hialinización. Además, otros factores como infecciones de la vía aérea superior, alergia, reflujo gastroesofágico y determinados perfiles psicológicos como personalidad extrovertida se han visto relacionados con la aparición de esta patología (Karkos & McCormick, 2009).

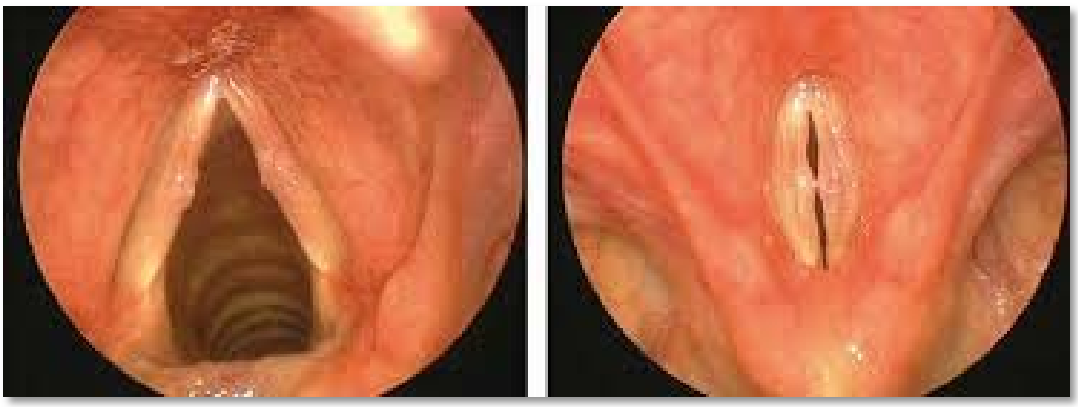

Figura 3.1. Vista de nódulos vocales en cuerdas vocales en abducción (izquierda) y aducción (derecha) (Núñez & Fernández, 2013).

A lo largo de los años, diferentes autores han apuntado a la relación entre la incidencia de nódulos y el género femenino. Ford (1994) afirmaba que son comunes en mujeres a partir de la segunda década de la vida y Perelló (1977) recoge que hasta el 81,7% de los nódulos aparecen en mujeres. Dependiendo del estadio en el que se encuentren los nódulos, se pueden observar distintas modificaciones en el epitelio de la cuerda vocal. Los nódulos agudos presentan un estroma edematoso, sin cambios epiteliales destacados. Por el contrario, en los nódulos más evolucionados, se puede apreciar hiperqueratosis con engrosamiento de la membrana basal.

Este aumento de masa de la cuerda vocal se traduce en una disminución de la Fo, con dificultad para la emisión de tonos agudos y perturbación de la vibración, convirtiéndose en aperiódica. Esto produce un aumento de los parámetros Jitter y Shimmer (Cobeta, 2008). El tratamiento de los nódulos vocales está basado en métodos conservadores que incluyen, reposo vocal, terapia vocal y tratamiento de otras enfermedades que pudieran estar actuando como factores desencadenantes o favorecedores (Núñez & Fernández, 2013).

El pólipo vocal es una lesión pedunculada o sésil, generalmente unilateral y vascularizada, en la que se produce una disrupción de la membrana basal. Su localización más frecuente es el tercio anterior de la cuerda vocal (Kleinsasser, 1986). Tiene mayor prevalencia en el género masculino. Su tamaño puede ser variable, aunque generalmente es mayor al de los nódulos y su aspecto es angiomatoso, edematoso e inflamatorio. Puede variar entre un color rojizo de apariencia sanguinolenta cuando su origen es hemorrágico, o blanquecino e incluso traslúcido cuando no lo es (véase la Figura 3.2). Su etiología está relacionada al igual que en los nódulos con el mal uso y/o abuso de la voz, además de otros factores favorecedores como el consumo de alcohol, esfuerzos con glotis cerrada, fonotraumatismos o alteraciones en la vía respiratoria superior.

Figura 3.2. Visa de pólipos vocales (Núñez & Fernández, 2013).

La clínica del pólipo vocal es variable, distinguiéndose disfonía brusca e irregular, disminución del rango tonal, timbre variable o bitonal, fonastenia, o sensación de cuerpo extraño con carraspeo habitual. Existe una disminución de la frecuencia fundamental, con afectación de los parámetros Jitter y shimer, así como los relacionados con el ruido. El tratamiento de los pólipos incluye cirugía en aquellos casos en los que existe una falta de respuesta significativa a los tratamientos conservadores (terapia e higiene vocal). También en aquellos casos en los que el pólipo se asocia a una variz, ya que existe un alto riesgo de hemorragia (Núñez & Fernández, 2013).

El pseudoquiste seroso es una lesión exudativa del espacio de Reinke de carácter unilateral que se sitúa en el borde libre de los pliegues vocales. Se caracteriza por presentar ligera hinchazón local (Koufman & Belafsky, 2001).

En ocasiones, suele considerarse a esta patología como una forma clínica de un nódulo (véase la Figura 3.3). Su etiología es similar a la de nódulos y pólipos en cuanto a la participación de factores relacionados con el abuso vocal. Además, se ha visto la influencia de otras entidades como alergias, infecciones, contaminación ambiental o trastornos endocrinológicos (Karkos & McCormick, 2009) en el desarrollo de edema localizado. La modificación estructural que pueden producir estas situaciones en la cuerda vocal puede desembocar en un pseudoquiste.

Figura 3.3. Vista de pseudoquiste seroso (Núñez & Fernández, 2013)

El edema de Reinke es una lesión que afecta a la capa superficial de la cuerda vocal. Su presentación es bilateral, existiendo una distensión de la parte superficial de la lámina propia que deforma la superficie cordal (véase la Figura 3.4). Esto es debido a la presencia de una sustancia inflamatoria gelatinosa que expande el espacio de Reinke desde la comisura anterior hasta el proceso vocal del Aritenoides. Puede acompañarse de congestión vascular y estasis venoso en la región afectada. Su etiología está ligada al hábito tabáquico, pero existen otros factores favorecedores como son el reflujo gastroesofágico, los trastornos tiroideos o el mal uso o abuso vocal.

Figura 3.4. Vista de edema de Reinke (Bonilla, Pérez & Camacho, 2018).

El edema de Reinke presenta una clínica caracterizada por una disfonía progresiva con Fo disminuida, lo que le confiere un tono vocal disminuido.

Este hecho favorece que su identificación inicial sea más habitual en el sexo femenino, ya que los tonos graves son más tolerables en el masculino (Núñez & Fernández, 2013). Las fuerzas aerodinámicas y mioelásticas implicadas en la fonación se ven incrementadas en esta patología debido a la hiperfunción vocal que lleva aparejada. Esto conlleva una alteración en la estructura anatómica de los pliegues vocales, y un aumento del tamaño de las lesiones (Zeitels et al.1997). A su vez, esto produce un aumento de tejido por edema de la submucosa, hecho que acentúa aún más el estrés mecánico sufrido en las cuerdas vocales, cerrando así un círculo vicioso que es común a todas las lesiones exudativas del espacio de Reinke (Pickhard & Reiter, 2013). A partir de un modelo matemático y físico se demuestra que el fluido contenido en los vasos de las cuerdas vocales se desplaza hacia el punto de éstos donde mayor amplitud de vibración exista, siendo proporcional a la frecuencia y al cuadrado de la amplitud (Czerwonka et al., 2008).

El quiste subepitelial es una lesión ubicada en la lámina propia de la cuerda vocal, generalmente en la superficie craneal del tercio medio cordal (Fazio et al., 2014) (véase la Figura 3.5). Principalmente se dividen en dos tipos: quistes epidérmicos y quistes de retención mucosa. Los quistes epidérmicos se sitúan en las capas subepiteliales de la cuerda vocal. Están formados por una pared de epitelio estratificado y queratinizado (Martins et al., 2011) y un contenido caseoso producto de la descamación del epitelio, lo que les confiere un aspecto perlado. Su aparición se relaciona con el abuso vocal (Núñez & Fernández, 2013), existiendo autores que apuntan a un origen congénito por acumulación de residuos de epitelio del cuarto y sexto arco faríngeo (Fazio et al., 2014). Los quistes de retención mucosa se producen por la obstrucción de los conductos glandulares del pliegue vocal, contribuyendo a su formación factores como el abuso vocal, una infección respiratoria de vías altas o el reflujo gastroesofágico. Son más comunes en los adultos con altos requerimientos vocales, especialmente en el sexo femenino (Milutinović & Vasiljević, 1992), aunque también pueden aparecer en la infancia. Su localización puede variar respecto a los epidermoides, siendo frecuente encontrar una lesión contralateral por roce (nodular o hiperqueratósica).

Figura 3.5. Vista de quiste subepitelial (Chuang, Badía & Napolitano, 2017).

Por último, bajo la denominación lesiones vasculares de las cuerdas vocales se encuentran una serie de patologías cuyo origen es consecuencia de la dilatación o rotura de vasos sanguíneos a lo largo de la cubierta mucosa del pliegue vocal. Se desarrollan en la zona de mayor impacto vocal, produciendo una clínica variable que puede ir desde síntomas sutiles prolongados en el tiempo, como disfonía o fatiga vocal, hasta disfunciones importantes del cierre glótico con diplofonía. Entre estas lesiones se encuentran las varices, las hemorragias subepiteliales, las ectasias y los pólipos hemorrágicos (Núñez & Fernández, 2013; Cobeta et al., 2013c).

3.1.4. Características vibratorias de nódulos, pólipos y edema de Reinke

El análisis del comportamiento vibratorio de los pliegues vocales es fundamental para determinar las causas de una disfonía. La estructura anatómica individual del aparato fonador, y especialmente de las cuerdas vocales, determinará las características del patrón vibratorio. Por este motivo, la identificación estructural a través de técnicas de imagen de alta resolución en tiempo real se convierte en una herramienta fundamental, ya que permite determinar el comportamiento de cada una de las estructuras del pliegue vocal.

En este apartado se describen las características vibratorias de las lesiones exudativas del espacio de Reinke más relevantes y que son objeto de nuestro estudio. Estas lesiones comparten características estructurales, de comportamiento vibratorio y acústico que permiten establecer pautas de prevención, seguimiento y, en ocasiones, tratamientos similares.

La vibración de la onda mucosa de los pliegues vocales, tras el paso de aire por la glotis aducida, conforma un ciclo vocal que genera sonido a partir de movimientos de separación y aproximación de las cuerdas vocales. La señal sonora producida por la sucesión de ciclos tendrá una frecuencia que dependerá del número de veces que se produzca el contacto entre pliegues, y una intensidad que vendrá dada por la fuerza con que se produzca dicho contacto. La variación en la frecuencia de ciclos dependerá de diversos factores asociados, entre los que destaca la anatomía individual y el sexo. La voz del hombre vibra en un promedio de 100-150 ciclos por segundo, mientras que la de una mujer lo hace entre 200-270 ciclos, y la de un niño entre 220-280 ciclos (Cobeta et al., 2013b). Debido a la velocidad con la que se produce la emisión de un sonido vocal (ciclos/segundos), no es posible el visionado del fenómeno a partir de una cámara de visión normal, como es la del fibrolaringoscopio. El ojo humano puede percibir fenómenos dinámicos de hasta cinco o seis imágenes por segundo, pero los movimientos de la mucosa vocal precisarían el uso de una tecnología de visionado de alta velocidad. Para conseguir, por tanto, una imagen nítida en la retina humana, es preciso reducir la velocidad o detenerlo, hecho que se consigue a través de la ilusión óptica de la estroboscopia.

Las características propias de cada patología vocal hacen que la relación entre esta y su patrón vibratorio no sea constante. La localización, extensión, tamaño y las particularidades en la fonación de un paciente definirán en gran medida la vibración de la onda mucosa. A pesar de esto, existen patrones típicos de vibración que permiten la asociación a un determinado trastorno vocal.

El contenido en fibras de colágeno confiere a los nódulos vocales un aspecto blanquecino en la mayoría de los casos. Producen un edema localizado en la capa superficial de la lámina propia que a menudo interfiere en el cierre gótico, dando un aspecto de reloj de arena la glotis cerrada. A pesar de su localización anatómica, su tamaño no siempre guarda relación con el nivel de disfonía que provocan. La vibración de la onda mucosa es capaz de transmitirse por el pliegue vocal, aunque en el caso de los nódulos evolucionados puede estar muy enlentecida. Otro dato a considerar en este tipo de patrones vibratorios es la presencia de onda mucosa en el lugar de la lesión cuando en el nódulo predomina el edema, y la ausencia en ese punto cuando predomina la fibrosis (Cobeta et al., 2013b). La amplitud de vibración y la onda mucosa son normales en la valoración estroboscópica de nódulos con pequeña afectación del borde libre de la cuerda. Sin embargo, en nódulos firmes, con aspecto de masa, la amplitud de vibración y la onda mucosa estarán reducidas. Estos casos se corresponden con lesiones crónicas en las que predomina el abuso vocal intenso y continuado. La valoración del patrón vibratorio a través de técnica estroboscópica de este tipo de lesiones permite ajustar el diagnóstico (Spiegel et al., 1991) y, por tanto, el tratamiento. Los nódulos edematosos son lesiones que responden de forma eficaz a la terapia vocal y al cambio de hábitos fonatorios, mientras que los fibrosos suelen tener peores resultados ante este tratamiento.

Los pólipos vocales son lesiones que se desarrollan en el borde libre de la cuerda vocal. Según sus características estructurales y localización, pueden ser pediculados o sésiles, unilaterales o bilaterales, angiomatosos o blanquecinos. Dependiendo de la flexibilidad del pólipo, la onda mucosa puede estar aumentada o disminuida. Aquellos casos en los que predomine el edema en la lesión, el movimiento vibratorio de la mucosa puede estar aumentado en la cuerda afectada. Sin embargo, la onda mucosa estará disminuida en su amplitud y frecuencia (aperiódica) en los casos en los que la hemorragia, la proliferación colágena o la inflamación estén presentes en la lesión (Morente et al., 2001). Al igual que en el caso anterior, los pólipos producen un defecto del cierre glótico en forma de reloj de arena. Esta incapacidad de cierre permite que el aire se escape durante la fonación, disminuyendo la presión de aire espirado y, por tanto, la energía al inicio del ciclo vocal, lo que se traduce en una disminución de la vibración y tamaño de la onda mucosa.

El edema de Reinke es una patología que afecta a la capa superficial de la lámina propia del pliegue vocal. Se caracteriza por la presencia de un material inflamatorio gelatinoso en el espacio de Reinke. El tamaño de la lesión incidirá en la calidad y características del comportamiento vibratorio. En estadios iniciales, donde la afectación no implica pérdida de flexibilidad del pliegue vocal, la amplitud de la onda será mayor y los movimientos asimétricos con frecuencia aperiódica. El desplazamiento de movimientos vibratorios horizontales suele ser pequeño a pesar de que la onda esté aumentada.

Con el crecimiento de la lesión, las características de vibración varían, encontrándose amortiguación e incluso ausencia. Esto es así debido al efecto masa que la lesión puede provocar (Núñez & Fernández, 2013).

3.2. Uso profesional de la voz

La voz constituye la herramienta fundamental de trabajo en numerosas ocupaciones. Aquellas profesiones que por sus características precisen el uso de la voz durante periodos prolongados de tiempo, tendrán más posibilidades de producir alteraciones vocales que aquellas donde puedan existir periodos de descanso vocal o el uso vocal no sea relevante. Las implicaciones que los trastornos vocales pueden tener en diversos campos (social, laboral, económico, emocional…) acentúan la necesidad de prevenirlos y tratarlos en estadios iniciales (Farías, 2013).

A pesar de que las relaciones de causalidad entre la aparición de un trastorno vocal y el desempeño de una profesión no aparecen reflejadas en muchos datos epidemiológicos en la actualidad, existen investigaciones llevadas a cabo entre determinados grupos de trabajadores que sugieren que diferentes trastornos vocales están relacionados con la labor profesional desempeñada (Preciado et al., 2008; Russell et al., 1998). Zagoruiko (1982) definió el uso profesional de la voz como un sistema funcional entre el individuo y el entorno laboral que lo rodea que se establece cuando su actividad laboral precisa de la utilización de la voz como herramienta principal de trabajo. Esto supone la adaptación de las funciones fonatorias del individuo a las demandas requeridas por las circunstancias laborales y la interacción social.

La OIT considera a los docentes como el primer grupo de riesgo de padecer enfermedades profesionales de la voz (Behlau et al., 2012; Farías, 2013; Niebudek-Bogusz & Śliwińska-Kowalska, 2013). La OIT considera que la diversificación de actividades educativas y la variedad de espacios donde se imparte docencia son factores causantes de riesgo para la seguridad y salud en entornos educativos (Farías, 2013). Los problemas relacionados con la voz conllevan una clínica que abarca numerosos y diversos signos y síntomas. Algunas de las manifestaciones que refieren los docentes en las consultas de otorrinolaringología son la fatiga vocal, la sensación de cuerpo extraño en la garganta, el esfuerzo vocal o los cambios e incluso pérdida de tono y timbre. Estas circunstancias requieren un ajuste en las actividades que realizan en su jornada laboral, hecho que puede tener implicaciones en la calidad de la enseñanza y en el aprendizaje de los estudiantes (Roy et al., 2004). Núñez & Moreno (2013) clasifican los signos y síntomas vocales que refiere la población general en dos grupos. El primero incluye síntomas fonatorios como fatiga vocal, esfuerzo vocal, variación de la calidad de la voz tras la fonación, dificultad para proyectar la voz y pérdida del rango tonal para el canto.

En el segundo quedarían englobados síntomas faringolaríngeos como sensación de sequedad y dolor de garganta, carraspeo y gusto ácido o amargo. Los síntomas fonatorios se dan con mayor frecuencia en maestros y profesores, lo que hacer presumir que las demandas profesionales se centran en las estructuras laríngeas.

La Organización Mundial de la Salud y la OIT desde 1966 incluyen en sus recomendaciones de salud laboral la vigilancia de la salud de los docentes. A pesar de esto, no ha sido hasta finales del siglo XX cuando los trastornos vocales de los docentes han empezado a considerarse dentro de las enfermedades ocupacionales (Farías, 2013). Algunos autores han sugerido el término "voz ocupacional" para referirse a la salud y seguridad vocal del individuo que utiliza la voz como herramienta a través de la que desarrolla una profesión (teleoperadores, vendedores). Mientras que han definido "voz profesional" a aquella a partir de la que se desarrolla una profesión (Elhendi et al., 2012; Gañet, et al., 2007). Este hecho ha contribuido a que durante algún tiempo no existiera un consenso claro a la hora de definir los trastornos vocales como enfermedad profesional.

En España, el concepto de enfermedad profesional viene definido en el artículo 157 del texto refundido de la Ley General de la Seguridad Social aprobado por el Real Decreto Legislativo 8/2015 de 30 de octubre como *"*la contraída a consecuencia del trabajo ejecutado por cuenta ajena en las actividades que se especifiquen en el cuadro que se apruebe por las disposiciones de aplicación y desarrollo de esta Ley, y que esta proceda por la acción de elementos o sustancias que en dicho cuadro se indiquen para cada enfermedad profesional*".* Pero no es hasta el año 2006 cuando las enfermedades de la voz, en concreto los nódulos vocales, se reconocen en el cuadro de enfermedades profesionales con el código 2L0101 (RD 1299/2006 de 10 de noviembre) dentro del grupo "Enfermedades producidas por agentes físicos", precisando como actividades capaces de producirlos aquellas "actividades en las que se precise un uso mantenido y continuo de la voz, como son profesores, cantantes, actores, teleoperadores, locutores".

3.3. Salud vocal del docente

Las demandas vocales a las que se enfrenta un profesor o maestro en el desarrollo de su actividad laboral diaria suponen una adaptación continua de su función fonatoria. Modificaciones en la intensidad, el tono y/o el timbre son comunes entre el colectivo de docentes. Las condiciones del entorno en el que se desarrolla la actividad docente suponen uno de los aspectos importantes a tener en cuenta a la hora de la valoración de su salud vocal y de la corrección de conductas y hábitos vocales poco saludables. La prevalencia de los trastornos de la voz en los docentes es difícil de establecer, ya que muchos de los profesionales dedicados a la enseñanza integran las alteraciones vocales como parte natural de su práctica laboral y no notifican enfermedad alguna. A pesar de esto, existen estudios que estiman que la prevalencia de lesiones vocales en la población general se sitúa entre el 6-15%, mientras que en los docentes llegaría hasta el 90% (Hazlett et al., 2011; Martins et al., 2014; Williams, 2003).

Van Thal (1961) definió los términos abuso vocal y mal uso vocal para distinguir dos situaciones vocales diferentes. La primera de ellas hace referencia al estado de hiperfunción laríngea que se produce como consecuencia del uso de la voz durante periodos prolongados de tiempo. La segunda se refiere al patrón habitual de fonación hiperfuncional.

Siguiendo esta base de razonamiento, Núñez y Moreno (2013) relacionan el abuso vocal con hábitos en los que los mecanismos vocales se utilizan de forma excesiva en conductas no fonatorias que generan o pueden generar traumatismos en los pliegues vocales, como son el carraspeo, la tos continua, la risa, el llanto o el hábito tabáquico. También entienden el mal uso vocal como aquella fonación excesiva o inadecuada, en la que se emplea un ataque vocal duro o un tono vocal inapropiado. El abuso vocal está relacionado con un déficit de higiene vocal e incluye conductas como (Prater & Swift, 1984):

- Gritar o chillar, donde se produce una hiperaducción y violenta vibración del borde libre de los pliegues vocales. Esto puede conllevar desde una ingurgitación vascular a un hematoma submucoso.
- Vocalizaciones forzadas, que suelen producirse con tonos agudos e intensidad elevadas con la laringe en hiperaducción. Por ejemplo, la imitación de sonidos de coches o cohetes conllevan la hiperaducción laríngea, así como el intento de fonación mientras se carga o empuja elementos pesados.
- Carraspear (o aclaramiento de la voz) y tos recurrente. Estas conductas suelen producirse cuando existe una deficiente hidratación de la mucosa laríngea.
- Ataque glótico duro. La aproximación de las cuerdas vocales antes del inicio de la espiración provoca un aumento en la presión subglótica, y el inicio de la fonación se produce de forma abrupta.
- Hablar en exceso o hablar cuando existen alteraciones de las cuerdas vocales.
- Inhalación de sustancias irritantes (gases, polvo, tiza, humo de tabaco…).

En el mal uso vocal además de un uso incorrecto del tono e intensidad vocal, intervienen factores como la personalidad (agresiva, competitiva…), el ambiente de trabajo (altos niveles de responsabilidad) o la carga laboral (jornadas prolongadas). La utilización de un tono elevado de voz suele estar relacionado con niveles altos de intensidad. Esto ocurre con frecuencia en lugares con altos niveles de ruido en los que el ruido ambiental induce al individuo a elevar entre 13-17 Hz la intensidad de su voz en una conversación con tono alto (80 dB), 14-38 Hz si el tono es muy fuerte (90 dB) y hasta 21-74 Hz si se utiliza una voz a gritos (Núñez & Moreno, 2013). La tensión muscular que se origina en situaciones de intensidad emocional también contribuye a la elevación del tono y la intensidad vocal.

3.3.1. Factores de riesgo vocal

El uso profesional de la voz se caracteriza por una carga vocal excesiva que aumenta el riesgo de padecer disfonía. Garrido et al. (2004) establecieron cinco factores para definir la aparición de carga vocal. El primero de ellos correspondería con el tiempo de fonación durante una jornada de trabajo, cuyo índice se corresponde con el porcentaje de tiempo medio de fonación. En el caso de los profesores universitarios, se utiliza el tiempo de fonación por hora de clase. El segundo sería el nivel de intensidad empleado en la fonación (intensidad alta). El tercero hace referencia al tono medio de voz hablada (tono inadecuado). El cuarto estaría representado por el nivel de ruido ambiental (superior a 55dB). El quinto sería el tiempo dedicado al descanso vocal, que suele ser limitado.

Pero la voz se ve influida también por otros factores como:

- Condiciones acústicas ambientales representadas principalmente por el ruido ambiental (Efecto Lombard) y el nivel de reverberación. El nivel de ruido en un aula silenciosa está alrededor de los 20-30 dB, por lo que la intensidad de la voz empleada suele ser de unos 50 dB. Cuando el aula está ocupada por estudiantes, el nivel de ruido puede aumentar hasta los 50-65 dB. Si el docente durante el desarrollo de una clase con un nivel de ruido normal debe elevar la intensidad de su voz 15 dB (Efecto Lombard), llegaría a utilizar un nivel total de intensidad que puede ir hasta los 70-80 dB (Farías, 2013). Esto conlleva una serie de adaptaciones laríngeas (fisiológicas y estructurales) que, mantenidas en el tiempo, pueden desencadenar una lesión en las cuerdas vocales. Por otro lado, la estructura de las aulas, frecuentemente formada por materiales resistentes que generan reverberación, hacen que su acústica no sea la adecuada en la mayoría de los casos. El tiempo de reverberación depende fundamentalmente de dos parámetros: la absorción acústica de los materiales que la forman y el volumen de la sala. El tiempo de reverberación será alto cuando, tras generarse un sonido, la fuente que lo inició cese y el sonido vaya atenuándose poco a poco. Por el contrario, será bajo, cuando se atenúe rápidamente. Cuanto mayor sea el tiempo de reverberación, peor se escuchará el mensaje (tiempo de reverberación recomendado en aulas escolares: 0,7-1 segundos (Martínez et al., 2016).
- Sistema vocal. En el que se incluye el marco tónico postural, respiración fónica, articulación ágil y precisa (habla inteligible), fonación adecuada (vocalización), oído (autorregulación vocal), sistema nervioso (regulación del sistema).
- Factores químicos como tiza, polvo, tierra, ácaros o gases industriales, que pueden producir sensación de picazón en la garganta o sensación de cuerpo extraño.
- Factores físicos. Entre los que se incluyen la temperatura, iluminación, ventilación o humedad.
- Factores biológicos y psicosociales que englobarían la edad, el sexo, estado de salud general, estrés, personalidad...
- Hábitos de alimentación y consumo de sustancias tóxicas como tabaco, alcohol, alimentos irritantes que facilitan el reflujo gastroesofágico, medicamentos….
- Factores relacionados con la organización del trabajo como horarios de los turnos, descansos, carga de trabajo, espacios de trabajo…

Existen situaciones en las que los profesores se ven sometidos a sobrecargas vocales que conllevan adaptaciones funcionales inadecuadas durante la fonación (hiperfunción). Este hecho supone, en la mayoría de las ocasiones, lesiones laríngeas estructurales de origen funcional (Castejón, 2014).

Una demanda fonatoria por encima del tono y volumen habitual induce un aumento de la tensión muscular faríngea, laríngea y respiratoria. Esta tensión cuando es leve y/o limitada en el tiempo, puede dar lugar a fatiga vocal.

Sin embargo, cuando es mantenida en el tiempo y su intensidad es elevada, produce desajustes del sistema pneumo-fono-articulatorio, desencadenando una serie de círculos viciosos: bucle de esfuerzo vocal (o círculo vicioso del sobreesfuerzo vocal), bucle de modificaciones de lubricación (necesidad de tragar saliva más a menudo, beber agua, tomar un caramelo…), bucle de modificaciones auditivas y bucle de modificaciones estructurales (postura, músculos…) (Preciado et al., 2005). El bucle del sobreesfuerzo vocal fue utilizado anteriormente por Le Huche (1982) para referirse a aquellas situaciones en las que la emisión de la voz se ve comprometida y es necesario un aumento de energía para su producción. Esta conducta compensatoria que en un principio resulta eficaz, tiende a agotarse, dando como resultado un rendimiento ineficaz, con una reducción de la producción vocal. Así pues, cuánto más costosa es la emisión de voz, más se fuerza, y cuánto más se fuerza, menos se emite. De esta forma, se puede decir que conductas de sobreesfuerzo vocal llevan consigo una disminución de la relación consumo de energía-eficacia. Este comportamiento de sobreesfuerzo vocal, en muchas ocasiones se acaba convirtiendo en una conducta habitual, conduciendo a distorsiones duraderas del mecanismo vocal como utilización del soplo torácico superior en el momento de la proyección vocal, desorganización de la mecánica de la laringe, pérdida de la actitud vertical del paciente (fundamental para proyectar la voz de forma adecuada y eficaz), o pérdida de sensaciones subjetivas de fatiga (por acomodamiento) (Le Huche, 1982).

Existen una serie de acontecimientos que pueden provocar que el individuo continúe aumentando su esfuerzo vocal proporcionalmente a la disminución de su rendimiento vocal. Estos son los denominados factores favorecedores y desencadenantes (Le Huche, 1994). Los primeros son el resultado de particularidades que conciernen al sujeto y/o a su modo de vida:

- Obligación socioprofesional de hablar o cantar (factor más importante).
- Algunas características psicológicas como perfeccionismo, temperamento, nerviosismo, ansiedad…
- Situaciones psicológicas difíciles relacionadas con conflictos personales, dificultades profesionales…
- Afecciones crónicas de la esfera de ORL, como sinusitis, rinitis, laringitis…
- Exposición al ruido, al polvo, emanaciones irritantes, aire acondicionado…
- Consumo de alcohol y tabaco.
- Presencia de una persona disfónica (disfonía por imitación).

Los factores desencadenantes están relacionados con aspectos de naturaleza orgánica o psicológica, entre los que se encuentran procesos patológicos de ORL (laringitis aguda, traumatismo laríngeo, intubación endotraqueal...), factores psicológicos (estrés, depresión…), debilitamiento general (enfermedades, cansancio), tos irritativa, procesos alérgicos, periodo premenstrual (por engrosamiento del epitelio de los pliegues vocales) y embarazo e intervención abdominal (por obstrucción del soplo abdominal).

3.3.2. Niveles de prevención

El desarrollo de una disfonía funcional derivada del desempeño de una profesión docente puede ser el resultado de un proceso en el que participan tres componentes: el profesor y su circunstancia individual (hábitos, personalidad, estilo de vida…), el contexto en el que desarrolla su trabajo (espacio físico, condiciones acústicas, nivel de ruido…) y la interacción del individuo con el contexto en que trabaja (tono e intensidad utilizado al hablar durante la clase, en el patio, en el despacho, por teléfono…) (Castejón, 2014). Existen autores que afirman que la instauración de una disfonía funcional por el ejercicio de la profesión docente es más frecuente en los cinco primeros años de actividad laboral (Simberg et al., 2006).

La Ley de Prevención de Riesgos Laborales en España (Ley 31/1995) obliga al empresario titular, para el caso de centros concertados y privados, y a la Administración Pública, para el caso de centros públicos, a evitar los riesgos derivados del trabajo, evaluar los riesgos que no puedan ser evitados y tomar las medidas preventivas oportunas en relación a los riesgos detectados, informando y formando a los trabajadores respecto a los riesgos que pueden sufrir en su puesto de trabajo y sobre las medidas preventivas a adoptar para evitarlos. Cada país adapta su marco normativo en materia de prevención de riesgos laborales, no existiendo un consenso común. Sin embargo, sí existen ciertas directrices internacionales en materia de vigilancia de la salud en el trabajo, como son la Directiva 89/391/CEE en Europa, y el Convenio 155 de la OIT, que deberán tenerse en cuenta. En el caso de las universidades, los servicios de prevención son los encargados de velar por la salud y seguridad de los docentes, promoviendo e implementando medidas de corrección de situaciones peligrosas para la salud vocal, y desarrollando programas de prevención, promoción y protección de la salud, siguiendo la normativa vigente a este respecto.

La prevención en el profesorado universitario debería ocuparse de dos aspectos importantes: la competencia (procedimientos para saber actuar en el aula) y la motivación (querer cuidar y mejorar la voz) (Castejón, 2013). Existen diferentes modelos de prevención de disfonías propuestos en la literatura.

Desde un enfoque tradicional, podemos encontrar un modelo médico, en el que la prevención es entendida desde un punto de vista individual, tratando aspectos relacionados con la técnica vocal y la higiene vocal, y desatendiendo factores personales y contextuales (el espacio de trabajo). Los modelos integradores dirigen sus actuaciones hacia una prevención de la disfonía en la que se consideran factores personales, contextuales y la interacción de ambos. Siguiendo este modelo, la Clasificación Internacional del Funcionamiento de la OMS en 2001 considera que para explicar la condición de salud de una persona que padece disfonía hay que tener en cuenta funciones y estructuras corporales, actividad y participación profesional, así como factores personales y ambientales. Los modelos basados en la ergonomía, por su parte, valoran más los contextos naturales en los que existe una interacción del profesor y el espacio del aula durante la impartición de una clase.

Por último, destacarían aquellos modelos ecológicos, naturalistas, en los que el esfuerzo preventivo y posteriormente rehabilitador se centra en el trabajo que se desempeña en el aula (Castejón, 2014).

Diferentes autores, que han desarrollado parte de su investigación en el campo de la prevención de disfonías en el ámbito docente, coinciden en la necesidad de seguimiento de las funciones vocales durante los 5 primeros años de labor docente, incidiendo en aquellos que hayan realizado alguna consulta médica por alteraciones en la voz (Preciado et al., 2005; Simberg et al., 2004; Simberg et al., 2006; Vilkman, 2004). Castejón (2014) propone acciones preventivas en el profesorado universitario basadas en tres niveles de formación. En el nivel conceptual y actitudinal, incluye una formación inicial y continuada en la que se traten aspectos relacionados con la fisiología de la voz y situaciones patológicas de la misma, las propiedades del sonido en cuanto a su transmisión y su relación con las características acústicas de las aulas, identificación de los signos y síntomas de sobrecarga vocal, sin olvidar las medidas de higiene vocal. Desde un nivel procedimental, el profesor recibirá formación y entrenamiento en técnica vocal (posición y tono muscular, tipo de respiración, dicción, tono, intensidad y velocidad, proyección y prosodia). Además, se trabajará la identificación personal de las debilidades y fortalezas en cuanto a la emisión vocal. Por último, y desde un nivel ecológico, se fomentará la aplicación en el aula de las técnicas aprendidas en los niveles anteriores.

3.4. Programa de prevención de disfonías en docentes

El programa de prevención que se propone en esta sección trata de dar respuesta a la necesidad de intervenciones tempranas y continuadas a través de distintos niveles de prevención con ayuda de un sistema automático basado en características acústicas. Aunque en esta tesis el programa de prevención se aplica en el contexto del profesorado de la Universidad de Extremadura, constituye una propuesta general para otros profesionales de la voz. El experimento realizado con docentes de la UEx sirve como validación de la propuesta desarrollada.

El programa consta de dos partes. En la primera, se realiza una clasificación de los profesores en tres grupos de riesgo (bajo, medio y alto) de padecer nódulos vocales u otras patologías laríngeas con características similares, en función de los resultados del análisis acústico realizado. En la segunda, se asignan distintos niveles de prevención (primaria, secundaria y terciaria) en función del grupo de riesgo al que pertenezcan, y se aplican medidas relacionadas con el nivel de prevención correspondiente.

3.4.1. Asignación de docentes a grupos de riesgo

La valoración objetiva de los trastornos vocales puede llevarse a cabo a través de distintos métodos. En el diagnóstico de la patología vocal están incluidos cuatro elementos de valoración que pueden aplicarse de forma ordenada.

El primero consistiría en una anamnesis, a continuación, una valoración acústica subjetiva (GRABS), seguidamente una exploración laríngea con imagen, y por último un análisis acústico (Zraick, 2011). El análisis acústico además de proporcionar información sobre la calidad de la voz aporta una información cuantitativa para el estudio de la función vocal. Teniendo en cuenta que la función vocal es un fenómeno que engloba distintos aspectos, el análisis acústico multidimensional de varios parámetros resulta más adecuado para la descripción de voces patológicas (Karnell et al., 1995; Titze & Winholtz, 1993).

Las patologías vocales son frecuentes en los profesionales de la voz y, especialmente, en los docentes. La evaluación de los riesgos vocales derivados de la labor docente, generalmente se realiza en función de factores como la antigüedad en el puesto, la reverberación en el aula, el tiempo de clase y el tiempo de descanso vocal. La falta de programas preventivos formales influye negativamente en la salud vocal del personal docente. Motivado por esta falta de programas preventivos y por la dificultad de llevarlos a cabo con un coste razonable, se diseñó este programa de atención preventiva de la voz basado en señales acústicas. Este programa se ha aplicado en la Unidad de Salud Laboral del Servicio de Prevención de Riesgos Laborales de la Universidad de Extremadura.

La asignación de riesgos se basa en la información obtenida a partir de los parámetros acústicos de sujetos clasificados previamente en dos grupos: grupo control (sanos), y grupo de individuos con padecimiento de nódulos vocales. Se entrenó un modelo de clasificación utilizando las características acústicas de estos dos grupos para, más tarde, aplicar los parámetros obtenidos a los parámetros vocales de cada profesor, proporcionando su probabilidad de sufrir nódulos vocales u otras patologías con un lenguaje disfónico similar. De acuerdo con estas probabilidades, los profesores se dividieron en tres grupos de riesgo: bajo, medio y alto. Los profesores con una probabilidad inferior a 0,5 fueron asignados al grupo de riesgo bajo, entre 0,5 y 0,9 al grupo de riesgo medio, y por encima de 0,9 al grupo de riesgo alto. Estos puntos de corte se eligieron en base a experimentos previos realizados en el Programa de Vigilancia de la Salud de los Trabajadores de la Universidad de Extremadura.

Este procedimiento basado en características acústicas proporciona un método de clasificación no invasivo y de bajo coste y ayuda al diagnóstico clínico a través de parámetros objetivos.

3.4.2. Asignación de niveles de prevención

Perelló y Salvá (1980) definen disfonía como la pérdida del timbre normal de la voz por trastorno funcional o alteración orgánica de la laringe. A la alteración del timbre se añade la de las otras cualidades de la voz como son intensidad y tono.

Una pérdida persistente de estas cualidades puede suponer, desde una disfonía funcional (frecuente entre el profesorado) producida por un trabajo excesivo y mal ejecutado, hasta una lesión grave (Le Huche & Allalí, 2003; Morrison & Rammage, 1996).

Algunos autores consideran que, tras 15 días con disfonía, un profesor debería someterse a una valoración clínica por un especialista de ORL que incluyera estudio con fibrolaringoscopio y estroboscopio (Garrido et al., 2004).

Al ser los nódulos vocales la única patología vocal contemplada en el cuadro de enfermedades profesionales recogidas en el RD 1299/2006, el objetivo de realización del presente trabajo se ha centrado en desarrollar un programa de educación para la salud adaptado a diferentes niveles de riesgo de padecer dicha patología o aquellas que tuvieran características acústicas similares (lesiones del espacio de Reinke, explicadas anteriormente en este capítulo).

Las acciones preventivas que se proponen se enmarcan en un modelo conceptual que explica la situación de un profesor con disfonía o riesgo de disfonía. A partir del modelo seleccionado se plantean propuestas diferentes de prevención. Se definen grupos homogéneos de riesgo según las tareas, ocupación o actividad de los trabajadores que los integran, teniendo en cuenta que cuanto mayor sea la homogeneidad de esos grupos, mayor será la especificidad conseguida. Una vez clasificados los trabajadores, se asignan los niveles de prevención. El objetivo de la prevención primaria consiste en realizar actuaciones para evitar o cuando menos reducir la exposición del individuo frente a causas conocidas y evitables, y así evitar el inicio de la enfermedad. El objetivo de la prevención secundaria es modificar la historia natural de la enfermedad por medio de la detección y tratamiento oportuno en la fase preclínica. De ese modo, se podrá prevenir el desarrollo de la enfermedad una vez instaurada. En la prevención terciaria el objetivo es evitar una mayor pérdida de función en el caso de patologías diagnosticadas. Las medidas recogidas en este último nivel de prevención irán encaminadas hacia la rehabilitación y reeducación en hábitos vocales saludables.

A continuación, se muestran las medidas propuestas en el primer nivel de prevención de nuestro programa para profesores de la Universidad de Extremadura. Están destinadas a aquellos trabajadores que hayan sido asignados al grupo de riesgo bajo. En este nivel primario, las actuaciones se centran en la higiene vocal (ambiente, persona y voz) y la técnica vocal.

- Medidas de higiene vocal: el contexto laboral. Se estudiará el contexto laboral de cada persona, vigilando aspectos como la humedad, la exposición a olores fuertes, polvo, aire acondicionado y calefacción, acústica, temperatura, corrientes de aire o ruido. Si hubiera elementos motivadores de riesgo vocal, se elaborará el correspondiente informe, y se dirigirá al centro donde se imparte la docencia. A su vez, se enseñarán actuaciones individuales que puedan disminuir o paliar el riesgo vocal a la espera de otras posibles intervenciones.
- Medidas de higiene vocal: la persona. Desde el punto de vista de la persona, se trabajarán aspectos relacionados con hábitos de vida. Se recomendará:
    - No fumar.
    - Tener patrones de sueño adecuados, dedicando al descanso el tiempo suficiente para no sufrir hipotonía por cansancio que puede dar lugar a una voz más grave y aérea.

  - Evitar o cuándo menos reducir el consumo de cafeína, teína o alcohol, sustancias que producen una pérdida de agua corporal y afectarán, por tanto, a la hidratación de las cuerdas vocales.
  - Mantener un equilibrio nutricional adecuado, evitando digestiones pesadas y favoreciendo el movimiento diafragmático. Restringir o evitar el consumo de bebidas carbonatadas, picantes o muy condimentadas que favorecen el reflujo gastroesofágico. Evitar bebidas y comidas demasiado frías o calientes que pueden alterar la mucosa faríngea que forma parte del sistema de resonancia de la voz.
  - Mantener una postura correcta a la hora de hablar. La proyección de la voz precisa de la utilización de la respiración diafragmática. Las compensaciones musculares pueden afectar a los músculos laríngeos y cervicales.
  - Consultar con un especialista, sin dejar pasar el tiempo, en el momento que se aprecien síntomas vocales o laríngeos.
- Medidas de higiene vocal: la voz. Se incidirá en la importancia de la autorregulación vocal, tomando consciencia de cuál es el uso que se está haciendo de la voz. De esta forma, se recomendará:
  - No hablar en ambientes ruidosos. El sistema de retroalimentación auditivo se ve alterado cuando existe mucho ruido ambiental, de manera que tenemos que hacer un esfuerzo vocal para poder oírnos.
  - No gritar ni hablar gritando, ya que es uno de los comportamientos de esfuerzo que más afecta a la salud vocal.
  - No hablar en exceso, intentando realizar periodos de descanso vocal de 15-20 minutos varias veces al día. No se debería hablar más de 4 horas seguidas al día y se deberían descansar 2-3 minutos por cada hora de habla (Coll, 2013).
  - No carraspear, toser o estornudar de forma violenta, ya que son comportamientos traumáticos para los pliegues vocales.
  - Evitar los ataques vocales bruscos, ya que producen una excesiva presión subglótica y tensión muscular que suponen un comportamiento de esfuerzo importante.
  - Hidratarse de forma correcta y frecuente. La mucosa vocal precisa estar hidratada para no perder flexibilidad. Se evitará el uso de caramelos mentolados que resecan la mucosa. Se utilizarán humidificadores en caso de ambientes muy secos.
  - Evitar tensiones musculares en la zona de la cara, cuello y hombros. Se prestará atención a la posición de la lengua y los dientes, evitando la presión de los mismos en momentos de reposo y habla.
- Técnica vocal. El entrenamiento en técnica vocal supone la base para poder adecuar los requerimientos vocales de un docente a las numerosas situaciones laborales que se le pueden presentar en el desempeño de su labor. Se procurará:
  - Mejorar y/o corregir la postura corporal y el tono muscular a través de la verticalización del individuo (no hablar con la cabeza girada ni movilizando o levantando material pesado). Se potenciará la comunicación no verbal.

- Instaurar el modelo de respiración costo-diafragmática, fundamental en la utilización de la voz proyectada, potenciando el apoyo abdominal, la regulación del soplo fonatorio y el ajuste de la coordinación respiración-habla.
- Ejercitar y colocar los resonadores a través de ejercicios de dicción, resonancia, motricidad oral o articulación.
- Realizar ejercicios de calentamiento vocal siguiendo un orden o guion.
    - Respiración. Movilización y autopercepción del movimiento del diafragma a través de la inspiración y espiración.
    - Emisión de distintos fonemas para potenciar la resonancia (/m/ con los labios cerrados) y calentar los músculos vocales y resonadores a través de la vibración (/r/ (rr)).
    - Movilización de cabeza y cuello con ejercicios suaves (inclinación, círculos y lateralización).
    - Aplicación de suave manipulación en dirección cefálica de la zona anterior (laríngea) y posterior del cuello.
    - Movilización de la mandíbula (subir y bajar, movimientos laterales)
    - Movilización de los labios (fruncir los labios y sonreír) y la lengua (moverla arriba y abajo y de forma lateral, doblarla apoyando en los dientes superiores e inferiores).
- Autorregulación vocal y relajación psicomotriz. Aprender a controlar el nivel de tensión y energía resulta fundamental. Es importante saber localizar la tensión de músculos cervicales, hombros, mandíbula o pectorales, ya que forman parte del mecanismo fonatorio.

El objetivo de la prevención secundaria es modificar la historia natural de la enfermedad por medio de la detección y tratamiento oportuno en la fase preclínica. En el caso de nuestro programa, la prevención secundaria se aplicará a aquellos profesores que sean asignados al grupo de riesgo medio o alto de padecer una disfonía, antes de ser diagnosticados por un especialista. Las grabaciones acústicas llevadas a cabo en la primera parte del estudio servirán para la detección de patologías en estadios precoces. Las medidas a adoptar irán encaminadas a evitar el progreso de la enfermedad.

La terapia vocal está indicada en estos casos, y puede adaptarse a las características individuales de cada profesor. En 1993 Stemple clasifica la terapia vocal en 5 grupos: terapia vocal sintomática, psicológica, etiológica, fisiológica y ecléctica. Para nuestro trabajo, decidimos utilizar la terapia vocal sintomática y la fisiológica, ya que los aspectos recogidos en las demás, ya se trabajan en las estas. La primera de ellas se basa en la corrección de los síntomas vocales, respiratorios y de resonancia, para mejorar la patología vocal (Boone, 1983). El resultado es una modificación directa de los síntomas, lo que proporciona una gran motivación al paciente para seguir esa línea de trabajo.

La terapia vocal fisiológica (Colton et al., 2006) es la más reciente, e intenta modificar la actividad fisiológica inadecuada para la fonación. A través de la coordinación de los sistemas respiratorios, de resonancia y vocal, consigue mejorar el tono y la elasticidad de los músculos laríngeos, así como las características de la mucosa vocal. Los ejercicios de tracto vocal semiocluido que forman parte de esta terapia, consiguen mejorar el timbre vocal, modificando indirectamente la intensidad, y la frecuencia.

De esta forma, disminuye la fuerza de contacto de los pliegues vocales, favoreciendo una mejor emisión vocal (evitando hiperfonación e hipofonación), y un menor traumatismo.

El último nivel de actuación presentado en el programa corresponde a la prevención terciaria. Los profesores asignados al grupo de riesgo alto recibirán este tipo de atención (una vez valorados por un especialista). Las grabaciones de voz, al igual que en el nivel anterior, ayudarán a realizar un seguimiento de los lesionados, así como a evitar una mayor pérdida de la función vocal. Las medidas en este nivel están encaminadas hacia la reeducación vocal, las técnicas se solapan con las que se han venido utilizando en los niveles anteriores de prevención. El aprendizaje del uso correcto de la voz debe pasar desde el conocimiento de la técnica vocal hasta el empleo de la misma. Tanto el trabajo de entrenamiento vocal desde niveles iniciales de prevención, como el de rehabilitación proporcionado por la prevención terciaria, consta de diferentes pasos que se deben trabajar de manera simultánea y progresiva, atendiendo siempre a las características, necesidades y dificultades de cada paciente. La fonación incluye el trabajo de numerosos sistemas y órganos del cuerpo íntimamente vinculados, por lo que el trabajo sinérgico de cada uno de ellos dará como resultado una voz sana. Por este motivo la rehabilitación debe aplicarse de forma personalizada, incidiendo en aquellos aspectos que cada individuo precise. Coll (2013) incluye seis elementos como base del tratamiento rehabilitador: relajación (no aplicable en todos los casos como norma), rehabilitación de la función respiratoria, resonancia (uso correcto de los resonadores), emisión vocal, articulación y modulación.

La salud vocal del docente es finalmente una responsabilidad que implica a diferentes sectores, además del propio individuo. Por esto la prevención realizada por los equipos sanitarios en los servicios de prevención, no puede separarse de la que deben hacer las propias universidades como institución, a través de la implementación de medidas como:

- Medidas preventivas constructivas y/o de carácter técnico, aprovechando los avances técnicos para el diseño de construcciones y herramientas más eficientes. Es necesario diseñar aulas con niveles acústicos óptimos, sin reverberación excesiva, aisladas del ruido exterior y con sistemas de amplificación vocal.
  Además, que cuenten con pizarras electrónicas o cañones de video que eviten la utilización de tiza. También es necesario tener buena calidad del aire y condiciones adecuadas de temperatura y humedad.
- Medidas preventivas organizativas. Diseñadas como parte de una organización racional adecuada del trabajo, que permite distribuir las horas de enseñanza favoreciendo pausas. También se deben incluir estrategias para evitar el estrés.
- Medidas preventivas individuales. La capacitación mediante cursos específicos y campañas de concienciación sobre la importancia del cuidado de la voz.

Capítulo 4. Materiales y métodos

El presente capítulo está dividido en cuatro secciones en las que se abordarán aspectos relacionados con el diseño de la investigación, las variables estudiadas, los instrumentos utilizados para la adquisición de la información, y el método y análisis estadístico.

4.1. Diseño del estudio

Para la realización del presente trabajo de tesis doctoral se llevó a cabo un estudio epidemiológico observacional descriptivo transversal de proyección analítica, para poder clasificar a los sujetos estudiados en distintos niveles de riesgo (bajo, medio, alto) de padecer nódulos vocales u otras patologías de la voz con características acústicas similares a éstos. Al tratarse de un estudio transversal, en el que la recogida de datos se refiere a un momento puntual de tiempo, no se puede asumir un periodo de inducción y de latencia entre la exposición y el desenlace de la patología. Sin embargo, al estudiar factores de riesgo que no se alteran con el tiempo (por ejemplo la variable sexo), la secuencia temporal entre la causa y el efecto se respeta mejor, ya que es una variable presente desde el nacimiento, es decir, anterior al desenlace, en cuyo caso se parece más a un estudio de casos y controles. Lo mismo sucede con la variable referida a la actividad profesional (docentes), ya que esta está presente y se mantiene de manera estable en el tiempo.

La ambigüedad de la secuencia temporal entre la exposición y el desenlace de la alteración vocal es uno de los inconvenientes principales de los estudios transversales. A pesar de que en nuestro estudio existen variables cuantificadas temporalmente (número de horas de docencia a la semana, años de experiencia docente…), no siempre se puede tener la seguridad de que la presencia del factor de riesgo haya sido anterior a la aparición de la enfermedad si se observan y se miden ambos fenómenos al mismo tiempo.

En cuanto a la proyección analítica del estudio, éste podría aproximarse a un diseño de casos y controles, en el que la selección de los individuos está basada en la presencia (casos) o ausencia (controles) de la patología vocal a estudio. Se evaluaron distintas variables presentes en los sujetos con patología vocal compatible con nódulos vocales como: horas de docencia semanales, años de actividad docente, o presencia de hábito tabáquico. Con ellas se trató de estimar si existía una mayor prevalencia de exposición en los sujetos con patología vocal (casos) en comparación con los que no la tenían (controles). Se recoge la duración de los síntomas, aunque de manera poco consistente, ya que dicho dato es aportado de forma subjetiva por el sujeto.

4.1.1. Descripción de la muestra

La muestra total utilizada para el estudio está compuesta por 170 participantes. Se constituyeron tres grupos: un primer grupo de 25 sujetos con patología de nódulos vocales, un segundo grupo con 25 individuos sanos y un tercero con 120 participantes, cuya característica común es su condición de Personal Docente e Investigador (PDI) de la Universidad de Extremadura.

Los participantes del primer grupo se reclutaron de la consulta de Disfonía del Hospital San Pedro de Alcántara de Cáceres (HSPA). De los 25 sujetos, 21 son mujeres y 4 son hombres, con una edad media de 42,44 años (±10,87).

El segundo grupo muestral se conformó a partir del Personal de Administración y Servicios (PAS) de la Universidad de Extremadura. De los 25 participantes, 21 son mujeres y 4 son hombres, con una edad media de 42,84 años (± 9,88).

Por último, el tercer grupo incluido en el estudio se obtuvo a partir del PDI en activo de la Universidad de Extremadura. De los 120 individuos, 45 son mujeres y 75 son hombres, con una edad media de 48.91 años (±8,52), una experiencia laboral media como profesores de 20,77 años (±10,93) y una media de horas de docencia semanales de 6,45 (±2,37).

4.1.2. Criterios de inclusión y exclusión

El criterio general de inclusión utilizado para la realización de este estudio fue ser español nativo, y conocer y hacer uso habitual del idioma español. Esta condición resulta fundamental, ya que se precisa de la vocalización del fonema /a/ en el protocolo de la investigación, y esta, puede variar de un idioma a otro. La edad de los participantes debía situarse entre los 18 y los 65 años, y el carácter de participación en los tres grupos descritos debía ser voluntario.

Para la inclusión en el primer grupo, los sujetos debían tener un diagnóstico clínico de padecer nódulos vocales emitido por un especialista en otorrinolaringología (ORL). La valoración del sujeto para la emisión del diagnóstico debía consistir en una exploración global, en la que se incluyera evaluación objetiva de las cuerdas vocales a través de fibrolaringoscopio y estroboscopia, así como evaluación perceptual a partir de pruebas que recogieran la semiología vocal (GRABS, VHI, TMF, índice s/z). Quedarían excluidos de la muestra para la participación en el primer grupo, aquellos sujetos que no cumplieran el criterio general de inclusión, así como aquellos que tuvieran un diagnóstico clínico distinto al de nódulos vocales (edema de Reinke, disfonía funcional, quistes, pólipos).

Los integrantes del segundo grupo no debían padecer ningún tipo de alteración o patología vocal. Este hecho debía ser evaluado por un médico otorrinolaringólogo.

Quedarían excluidos de la muestra aquellos sujetos que tuvieran una dedicación complementaria a su trabajo principal que conllevara el uso profesional de la voz (cantantes, locutores/as, entrenadores/as), además de aquellos que no cumplieran con el criterio general de inclusión.

Por último, podría formar parte del tercer grupo, cualquier profesor de la Universidad de Extremadura que cumpliera el criterio general de inclusión, independientemente de cuál fuera su estado de salud vocal.

4.1.3. Recogida de datos

La recogida de datos del trabajo que se presenta tuvo lugar en dos escenarios distintos, la Universidad de Extremadura y el Hospital San Pedro de Alcántara de Cáceres. El proyecto de investigación fue informado favorablemente por la Comisión de Bioética y Bioseguridad de la Universidad de Extremadura, así como por el Comité de Bioética del Hospital San Pedro de Alcántara de Cáceres.

Los integrantes del primer grupo fueron valorados en la consulta de disfonía del HSPA de Cáceres, entre los meses de enero y junio de 2017. La recogida de datos y el protocolo de investigación se realizaron en dicha consulta.

En el caso de los participantes del segundo y tercer grupo, la evaluación y recogida de datos tuvo lugar en el Servicio de Prevención de Riesgos Laborales de la Universidad de Extremadura (UEx), entre los meses de enero y noviembre de 2017.

Atendiendo a los derechos de información sanitaria, intimidad y respeto a la autonomía del paciente descritos en la “Ley 41/2002, de 14 de noviembre, Básica Reguladora de la Autonomía del Paciente y Derechos y Obligaciones en Materia de Información y Documentación Clínica”, todo participante en esta investigación ha recibido información comprensible y adecuada proporcionada por los responsables del proyecto, y han firmado de forma voluntaria un consentimiento informado en el que expresan su conformidad con la inclusión en el estudio.

Se describe a continuación el procedimiento completo realizado para la recogida de datos para cada uno de los tres grupos integrados en el estudio.

- Grupo 1. Nódulos vocales. Se analizaron 25 sujetos remitidos a la consulta de disfonía. Los pacientes podrían llegar desde las consultas de atención especializada, de atención primaria, o incluso desde la misma unidad de ORL general. En la consulta de disfonía el sujeto es recibido por un equipo interdisciplinar compuesto por uno o dos especialistas en ORL, una logopeda y una enfermera. Se le realiza una anamnesis en la que se recogen sus antecedentes personales y familiares más destacados, y se procede a la valoración objetiva y subjetiva de medidas de calidad vocal. Dentro de las valoraciones realizadas por la logopeda se incluyen el TMF, los tiempos máximos de soplo fonatorio, el índice s/z, el ataque vocal, la postura, tipo de respiración utilizada, impostación, y articulación.

Al especialista en ORL le corresponde la evaluación perceptual de la voz, realizada a través de la escala de GRABS. Además, se entrega al paciente un cuestionario con preguntas sobre aspectos relacionados con hábitos de higiene vocal y el VHI-30. El resultado de la espirometría (prueba no invasiva para evaluar la capacidad pulmonar de una persona) es otro de los aspectos valorados. La mayor parte de los pacientes realizan esta prueba en las primeras consultas de disfonía. A través de un fibrolaringoscopio, y con un abordaje nasal, el especialista en ORL valora la cavidad nasal, bucal y laríngea. Con esta técnica, se obtiene una imagen en tiempo real de estructuras implicadas en la respiración y la fonación, aportando una primera impresión sobre posibles alteraciones que pudieran interferir en el acto fonatorio. A continuación, se realiza la exploración de los pliegues vocales con un videoestroboscopio. Esta prueba diagnóstica permite conocer la frecuencia fundamental del paciente, así como el comportamiento de la onda mucosa, dato fundamental para la emisión de un juicio clínico relacionado con la patología vocal. Tras la valoración completa del paciente, se procede al establecimiento del diagnóstico. En el caso de que la patología diagnosticada sea nódulos vocales, se informa al paciente de la posibilidad de participar en la investigación. Si da de forma voluntaria su consentimiento, pasa a la última fase de la evaluación: la valoración acústica de la voz. Antes de la realización del protocolo de grabación, llevado a cabo por la profesional de enfermería, el paciente debe contestar una serie de cuestiones relacionadas con aspectos sociodemográficos, de actividad laboral, enfermedad, y valoración subjetiva de la incapacidad de su voz. Este último aspecto se lleva a cabo a través del cuestionario VHI-10, una versión reducida y adaptada del cuestionario VHI (Rosen et al., 2004). A continuación, se explica al paciente las tareas de grabación que debe realizar. Se efectúan tres grabaciones de la fonación de la /a/ sostenida, en un tono e intensidad cómodos durante cinco segundos.

- Grupo 2. Sanos. Los participantes que forman parte del grupo de sujetos sanos se escogieron entre el personal de administración y servicios de la UEx, que realiza su reconocimiento médico anual en el Servicio de Prevención de Riesgos Laborales de dicha universidad. Para ser incluidos en el estudio, deben aceptar de forma voluntaria firmando un consentimiento informado. En el Servicio de Prevención, los trabajadores son recibidos y atendidos inicialmente en la consulta de enfermería. Allí se les realiza una evaluación de distintos parámetros relacionados con las constantes vitales, medidas antropométricas o actividad cardíaca, y las pruebas complementarias marcadas por el Protocolo de Vigilancia de la Salud específico del puesto de trabajo. Posteriormente pasan a la consulta médica, donde reciben la atención de un médico especialista en Medicina del Trabajo. Este valora los resultados de las pruebas de laboratorio previamente realizadas y procede a su exploración clínica por aparatos y sistemas, entre los que se encuentra la valoración de la salud vocal.

Una vez firmado el consentimiento informado, los pacientes rellenan el mismo cuestionario que el primer grupo, en el que se incluye la prueba VHI-10 (Rosen et al., 2004). Posteriormente, se explica e inicia el protocolo de grabación, que es el mismo que en los grupos anteriores.

- Grupo 3. Profesores Universidad de Extremadura. Los sujetos que conforman este tercer grupo son elegidos entre el personal PDI de la UEx que ha realizado su reconocimiento médico anual en el Servicio de Prevención de dicha Universidad, y aceptan voluntariamente mediante la firma de un consentimiento informado su participación en la investigación. El procedimiento es el mismo que se ha descrito detalladamente para el grupo de individuos sanos. Los integrantes del tercer grupo pueden padecer o no patología o alteración vocal. Todos serán incluidos en este grupo, ya que el objetivo del estudio presentado en esta tesis doctoral es la clasificación del PDI de la Universidad de Extremadura en distintos niveles de riesgo de padecer nódulos vocales o trastornos compatibles.

4.2. Variables

Una vez recogida la información, el siguiente paso es generar el fichero de datos en el que se incluyen tanto los datos procedentes del cuestionario como de la extracción de características acústicas. Las variables acústicas se extrajeron utilizando un programa realizado en Matlab (Gamero, 2013). El número de variables con las que se ha trabajado suman un total de 37, y se incluyen según su naturaleza en distintos grupos que se definen a continuación.

- Variables sociodemográficas:

  1. Edad. Variable cuantitativa que establece la edad en años del paciente cuando se realizó el estudio.
  2. Sexo. Variable cualitativa que recoge dos opciones, hombre y mujer.
  3. Estado. Variable cualitativa que aporta información sobre el grupo de estudio al que pertenece el sujeto; recoge tres opciones, PDI, sano, paciente ORL.
  4. Patología. Variable cualitativa que indica si el sujeto padece una patología vocal; recoge dos opciones, Sí/No.

- Variables respecto a la actividad laboral que desempeña:

  5. Tiempo de docente. Variable cuantitativa que establece la experiencia docente en años del sujeto cuando se realizó el estudio.
  6. Clases semanales. Variable cuantitativa que informa sobre el número de clases a la semana que imparte el sujeto.

- Variables sobre antecedentes personales y hábitos del paciente. Variables cualitativas y cuantitativas que aportan información sobre posibles factores etiológicos que pueden iniciar o agravar una patología vocal.

  7. Fumador. Variable cualitativa que recoge la condición de fumador o no del sujeto; recoge dos opciones Sí/No.

  8. Tiempo de fumador. Variable cuantitativa expresada en años, que recoge el tiempo que el sujeto lleva fumando.

  9. Exfumador. Variable cualitativa que describe la condición de exfumador del sujeto; recoge dos opciones Sí/No.

  10. Tiempo de fumador. Variable cuantitativa que recoge en años, el tiempo que el sujeto ha estado fumando antes de dejarlo.

  11. Tiempo de exfumador. Variable cuantitativa que establece en años, el tiempo que el sujeto lleva sin fumar.

  12. Enfermedad neurológica o mandibular. Variable cualitativa que aporta información sobre los problemas relacionados con la articulación témporo-mandibular (implicada en la fonación) y/o alteraciones neurológicas. Recoge dos opciones Sí/No; en caso afirmativo, requiere una concreción del problema.

  13. Medicación habitual. Variable cualitativa que establece si el sujeto toma medicación de forma habitual; recoge dos opciones Sí/No. En caso afirmativo, requiere una concreción del tipo de medicación utilizada.

  14. Enfermedad crónica de la voz. Variable cualitativa que establece si el sujeto padece alguna alteración o patología crónica de la voz. Recoge dos opciones Sí/No. En caso afirmativo, requiere una concreción del problema.

- Variables sobre la historia de la enfermedad percibida. Variables cualitativas que informan sobre las características de la enfermedad percibida por el sujeto.

  15. Enfermedad actual de la voz. Variable cualitativa que informa sobre el estado actual de la voz de un sujeto; recoge dos opciones Sí/No.

  16. Problema actual con la voz. Variable cualitativa que aporta información sobre si el sujeto refiere problemas con la voz en el momento en el que rellena el cuestionario; recoge dos opciones Sí/No.

- Variable respecto al índice de incapacidad vocal percibida.

  17. VHI-10. Variable cuantitativa que mide el índice de incapacidad vocal global percibido por el sujeto en la escala reducida VHI.

- Variables referentes a los parámetros acústicos.

18. Jitter absoluto. Variable cuantitativa que mide la variación entre ciclos de la frecuencia fundamental de la muestra analizada, y se expresa en microsegundos (µs).
19. Jitter relativo. Variable cuantitativa que mide la variabilidad relativa ciclo a ciclo de la frecuencia fundamental. Es un indicador de la estabilidad de la Fo. Se expresa en porcentaje (%).
20. RAP. Variable cuantitativa que mide la perturbación de la Fo, promediando la diferencia de tres periodos consecutivos.
21. PPQ. Variable cuantitativa que mide la variabilidad de la frecuencia fundamental en grupos de cinco ciclos.
22. Shimmer local. Variable cuantitativa que mide la variabilidad entre periodos consecutivos de la amplitud (pico a pico) de la onda de voz analizada. Se expresa en decibelios (dB).
23. Shimmer relativo. Variable cuantitativa que determina la variabilidad entre ciclos consecutivos de la amplitud. Se expresa en porcentaje.
24. APQ. Variable cuantitativa que analiza la variabilidad relativa de la amplitud de la onda analizada en grupos de tres en tres, cinco en cinco u once en once periodos.
25. sAPQ. Variable cuantitativa que evalúa la variabilidad relativa de la amplitud en ciclos de 55 en 55.
26. NHR. Variable cuantitativa que determina el cociente entre el componente periódico (armónicos) y la energía aperiódica (ruido) de una señal acústica analizada, permitiendo realizar una evaluación general del ruido de dicha señal.
27. NNE. Variable cuantitativa que muestra los valores transformados de la intensidad de ruido encontrados en una muestra analizada, en valores de una distribución normal.
28. SNR. Variable cuantitativa que muestra la relación entre la energía de la señal y la energía de ruido.
29. GNE. Variable cuantitativa que mide la excitación vocal debido a la vibración de los pliegues vocales, frente a la excitación producida por turbulencias de aire espirado.
30. PermutEn. Variable cuantitativa que mide la entropía.
31. FuzzyEn. Variable cuantitativa de medición de entropía difusa.
32. SampleEn. Variable cuantitativa de medición de entropía muestral.
33. D2. Variable cuantitativa que caracteriza la distribución de puntos del atractor en el espacio de fase. Es una medida de complejidad de un sistema no lineal.

34. RPDE. Variable cuantitativa de medición de desviaciones con respecto a un comportamiento periódico perfecto.
35. DFA. Variable cuantitativa que mide las fluctuaciones sin tendencia de la señal acústica.
36. PPE. Variable cuantitativa que mide la entropía del tono.

4.3. Material técnico

Para la valoración funcional y estructural de las cuerdas vocales se ha utilizado un equipo de fibrolaringoscopio OLYMPUS Visera Élite OTV-S190 y CLV-S190, y estroboscopio KARL STORZ-ENDOSCOPE Pulsar 20140020 (véase la Figura 4.1).

Figura 4.1. Equipo de fibrolaringoscopio y estroboscopio.

La grabación de las distintas tareas vocales se llevó a cabo en dos espacios físicos distintos. Los integrantes del primer grupo (sujetos con patología vocal) fueron grabados en la unidad de voz del HSPA, en una habitación no insonorizada, pero sí aislada. Los participantes asignados al segundo y tercer grupo fueron grabados también en una habitación aislada no insonorizada integrada en la estructura física del Servicio de Prevención de la UEx.

La señal de la voz de los participantes en los distintos grupos fue grabada a través de un micrófono tipo cardioide AKG 520 (véase la Figura 4.2) conectado a un ordenador portátil con una tarjeta de sonido externa tipo TASCAM US-322 (véase la Figura 4.3). El software de grabación utilizado ha sido Audacity en su versión 2.0.5, utilizando una tasa de muestreo de 44.1KHz y una resolución de 16 bits.

Figura 4.2. Micrófono Cardioide AKG-520.

Figura 4.3. Tarjeta de sonido TASCAM US-322.

4.4. Métodos estadísticos

El procesamiento de señal fue implementado en Matlab, integrando los algoritmos de extracción en una interfaz gráfica con un módulo de análisis estadístico y otro de clasificación. De las variables analizadas, finalmente fueron diez las características acústicas que resultaron más efectivas para el propósito de la investigación. Tres de las características fueron lineales: Jitter, Shimmer y GNE; y siete no lineales: PermutEn, SampEn, FuzzyEn, D2, RPDE, DFA y PPE. Las características no lineales son de particular relevancia para la práctica clínica (Jiang et al., 2009). Las voces patológicas con disfonía severa son las que cuentan con más probabilidades de presentar fenómenos altamente no lineales y aleatorios, mientras que las voces saludables son las que se encuentran más cerca del modelo de filtro de fuente lineal. La variable sexo también fue incluida en el análisis, ya que es uno de los factores influyentes en el desarrollo de la patología vocal.

La anatomía de la laringe femenina (más pequeña que la de los hombres) y sus características estructurales, hacen que el uso de esta sea distinto en un género u otro, siendo necesaria una tensión mayor a la hora de la fonación en el caso de las mujeres, hecho que se traduce en una mayor tendencia a padecer patologías vocales.

Cada una de las medidas de análisis acústico utilizadas fueron promediadas para cada una de las tres repeticiones de /a/ sostenida, generando una matriz de 170 filas (una por cada sujeto) y 11 columnas (una por cada una de las características utilizadas más el sexo). Una vez extraídas las características, se implementó un procedimiento de selección de variables para elegir el subconjunto óptimo de predictores según las características de los grupos de análisis primero y segundo (Al-Nasheri et al., 2017).

La dimensionalidad del espacio de entrada se redujo según el principio metodológico de Occam o principio de parsimonia, según el cual, cuando dos teorías a igualdad de condiciones tienen como resultado las mismas consecuencias, debe darse preferencia al modelo más sencillo (Tsanas & Gómez-Vilda, 2013). En nuestro caso, el modelo más sencillo es aquel que cuenta con menos características, y mantiene la misma capacidad predictiva. Esto evita predictores innecesarios o redundantes que pueden agregar ruido, al mismo tiempo que reduce la carga computacional. Los datos obtenidos de la extracción de características y las encuestas fueron analizados con el software estadístico R.

Inicialmente se realizó un análisis descriptivo exploratorio de los datos. Se calcularon las medias, desviaciones estándar, diagramas de cajas y bigotes, y tablas de contingencia. A continuación, se realizaron pruebas inferenciales, considerándose estadísticamente significativas cuando los p-valores fueron inferiores a 0,05. Cuando las condiciones de aplicabilidad se pudieron asumir, el *t* test para variables independientes y el Análisis de la Varianza (ANOVA) se utilizaron para la comparación de medias. Cuando no se cumplieron las condiciones necesarias para utilizar contrastes paramétricos, se emplearon los tests de Mann-Whitney-Wilcoxon y Kruskal-Wallis.

Los contrastes sobre las medias de tres grupos se analizaron con los tests de Kruskal-Wallis o ANOVA, y en aquellos casos que resultaron significativos, se aplicaron las pruebas de Mann-Whitney penalizadas o de Bonferroni para establecer comparaciones por pares.

Para las variables cualitativas se empleó la prueba exacta de Fisher (Al-Nasheri et al., 2017) para analizar si existían asociaciones estadísticas. Para medir la intensidad de las asociaciones estadísticas entre variables cualitativas (cuando los test de homogeneidad resultaron significativos), se utilizó el índice *V* de Cramer.

Para estimar la probabilidad de padecer nódulos vocales u otras patologías con características acústicas similares, se utilizaron clasificadores de *machine learning* que consideraron modelos individuales y conjuntos (conocidos como *Ensemble Modelling*). Estos clasificadores fueron:

- *K- Nearest Neighbor (K-nn)*. Método de clasificación, que clasifica cada elemento en el grupo mayoritario de sus k vecinos más cercanos.

- *Random Forest (RF)*. Método de clasificación que segmenta el espacio de los predictores en regiones simples para el estudio de las interacciones de datos aleatorios con igual distribución. Utiliza una serie de árboles de decisión, con el fin de mejorar la tasa de clasificación. Técnica de gran eficacia en ciencias biomédicas para detectar lesiones físicas u orgánicas.
- *Vector Generalized Linear Model Adjoint Categories* (*VGLMs*). Modelo lineal generalizado, que permite utilizar variables de respuesta fuera de la familia exponencial clásica y para más de un parámetro.
- *Elastic-Net Regularized Generalized Linear Model.* Método de regresión logística que combina linealmente las penalizaciones de modelos Lasso y Ridge. En nuestro caso utiliza el algoritmo de optimización *cyclical coordinate descent*.
- *Averaged Neural Network.* Método estadístico utilizado para reconocer patrones establecidos, que permite clasificar vectores de entrada, haciendo un promedio del resultado de varias redes neuronales que se ejecutan al mismo tiempo.
- *Least Squares Support Vector Machine.* Conjunto de métodos de aprendizaje supervisados relacionados, que utiliza funciones lineales para resolver el SVM en lugar de funciones convexas cuadráticas. En este caso se utilizó un kernel polinómico.
- *Bagged Classification and Regression Trees.* Método no paramétrico predictivo, en el que a través de múltiples árboles de decisión se realiza un remuestreo de los datos, para hallar una predicción de consenso.
- *Boosted Logistic Regression.* Método estadístico multivariante de clasificación, cuyo objetivo es dar mayor relevancia en cada iteración a las observaciones mal clasificadas en pasos anteriores. Su objetivo es disminuir la varianza y el sesgo de las predicciones y prevenir el sobreajuste.
- *Bayesian Additive Regression Trees:* Modelo de clasificación donde cada árbol de decisión cuenta con una distribución a priori débil, y la clasificación se realiza después de un proceso iterativo de MCMC que genera muestras de la distribución a posteriori.
- *Conditional Inference Random Forest (RF).* Método estadístico en el que, tras realizar varios árboles de inferencia condicional, promedia los resultados para predecir una respuesta. Tiene la ventaja de no utilizar todas las variables en cada árbol. Con esto evita el problema de *overfitting* o sobreajuste.
- *Random Generalized Linear Model Ensemble.* Técnica utilizada para clasificación que incluye las ventajas de Random Forest (importancia de variables, nivel de acierto muy elevado), y cuenta con un modelo de selección de clasificadores lineales hacia adelante (interpretabilidad).

Para poder generalizar los resultados obtenidos con todos estos modelos, se realizó un proceso de validación cruzada con k-particiones. Esta técnica consiste en dividir aleatoriamente el conjunto de datos en k subconjuntos de igual tamaño. Se conserva un único subconjunto como datos de validación para probar el modelo, mientras que los k-1 subconjuntos restantes se usan como datos de entrenamiento. Este proceso se repite k veces, con cada uno de los k subconjuntos utilizados una vez, como datos de validación. Finalmente, los k resultados se promedian para producir medidas estadísticas generales del rendimiento del clasificador, como la proporción de acierto, el área bajo la curva de características operativas (AUC) del receptor, la sensibilidad y la especificidad. Ya que el estudio cuenta con 50 sujetos, y siendo k= 5, se formaron subconjuntos con 10 sujetos (cinco de VFN, y cinco de HC).

Se realizó una comparación de modelos y se seleccionó el modelo con la mejor proporción de acierto. Los parámetros del modelo, obtenidos con la información de los grupos de sanos y enfermos, se aplicaron a los predictores de los profesores, proporcionando, para cada uno de ellos, una probabilidad diferente de padecer nódulos vocales u otras patologías con características acústicas similares. Estas probabilidades se utilizaron para clasificar a los profesores en tres grupos de riesgo (bajo, medio y alto) en el programa de prevención de patologías vocales que se describió en el capítulo anterior.

Capítulo 5. Resultados experimentales

En este capítulo se presentan los resultados obtenidos del análisis del cuestionario realizado por los grupos de sujetos sanos, los que sufren nódulos y profesores universitarios. Así como el análisis de las características acústicas utilizadas en la investigación, también segmentados por grupo. Se aplica la metodología presentada en el capítulo de materiales y métodos, para, posteriormente, realizar la asignación a niveles de prevención de los profesores participantes en el estudio.

5.1. Análisis del cuestionario

En la Tabla 5.1 se muestran las frecuencias observadas de las variables cualitativas del cuestionario, divididas por grupo.

Variables		Grupo					
		Sano		Enfermo		Profesor	
		Recuento	%	Recuento	%	Recuento	%
Sexo	Hombre	4	16,0	4	16,0	75	62,5
	Mujer	21	84,0	21	84,0	45	37,5
Fumador	No	21	84,0	21	84,0	108	90,0
	Sí	4	16,0	4	16,0	12	10,0
Exfumador	No	14	73,7	12	48,0	77	66,4
	Sí	5	26,3	13	52,0	39	33,6
Enfermedad mandibular	No	25	100,0	20	80,0	116	98,3
	Sí	0	0	5	20,0	2	1,7
Problema voz	No	25	100,0	0	0	103	85,8
	Sí	0	0	25	100,0	17	14,2
Medicación	No	17	68,0	11	44,0	79	66,9
	Sí	8	32,0	14	56,0	39	33,1

Tabla 5.1. Tabla de contingencia para las variables cualitativas del cuestionario.

Se puede observar que en los grupos de enfermos y sanos la muestra se eligió para que fuera equilibrada respecto al sexo, por lo que el porcentaje de hombres es del 16% y el de mujeres es de un 84% en ambos casos. En el grupo de profesores el porcentaje de hombres es superior al de mujeres en una proporción de 62,5% frente al 37,5%. Esto se debe a la propia asistencia a las revisiones periódicas del Servicio de Prevención de Riesgos Laborales de la Universidad de Extremadura en las fechas que se realizó el estudio.

En cuanto al estado de fumador, el porcentaje de fumadores es igual entre sanos y enfermos (16%), mientras que entre los profesores este porcentaje es inferior en 6 puntos porcentuales. La proporción de exfumadores en el grupo de enfermos (52%) es prácticamente el doble que el de sanos (26,3%). El porcentaje de profesores exfumadores alcanza el 33,6%, porcentaje intermedio entre el de los sanos y el de los enfermos.

Los individuos pertenecientes al grupo de enfermos afirman tener problemas relacionados con trastornos neurológicos y/o mandibulares en un 20% y todos tienen problemas de la voz (100%). Los sanos no tienen ninguno de estos problemas. Los profesores se encuentran en un punto intermedio con algunos individuos con problemas neuromandibulares (1,7%) y con problemas de la voz (14,2%).

Todos los grupos de la muestra presentan porcentajes de medicación altos (superiores al 30%), siendo esta medicación variada e incluyendo en ciertos casos fármacos para problemas relacionados con la voz. Más de la mitad de los enfermos de la muestra está tomando alguna medicación (56%) frente a un 32% de los sanos. Entre los profesores, el 33,1% respondió afirmativamente, siendo un porcentaje similar al de los sanos.

La Tabla 5.2 representa algunos estadísticos descriptivos de las variables cuantitativas del cuestionario.

Grupo		Media	Mediana	Desv. típ.	Mínimo	Máximo
Sano	Edad	42,84	45,00	9,88	26	59
	Tiempo como fumador (años)	13,25	12,50	3,94	10	18
	VHI10	1,68	0,00	2,70	0	11
Enfermo	Edad	42,44	44,00	10,87	18	64
	Tiempo como fumador (años)	34,33	33,00	5,13	30	40
	VHI10	16,36	13,00	10,82	2	36
Profesor	Edad	48,91	49,50	8,52	28	67
	Tiempo como fumador (años)	22,08	21,00	11,90	1	40
	VHI10	3,53	2,00	5,02	0	28
	Tiempo docente (años)	20,77	22,00	10,93	0	43
	Clases por semana	6,45	6,00	2,37	3	12

Tabla 5.2. Estadísticos descriptivos para variables cuantitativas del cuestionario VHI.

La edad media del grupo de sanos es de 42,84 años (± 9,88), mientras que la del grupo de enfermos es 42,44 años (± 10,87). Tanto la posición central como la dispersión de las edades son similares. Esto ocurre porque el grupo de los sanos se ha elegido para aproximarse lo máximo al grupo de enfermos.

Sin embargo, los profesores han sido elegidos según han asistido a las revisiones periódicas en las fechas establecidas para el estudio.

La edad media de los profesores es superior en algo más de 6 años (48,91 ± 8,52). Las distribuciones de edad se muestran en los diagramas de cajas y bigotes de la Figura 5.1.

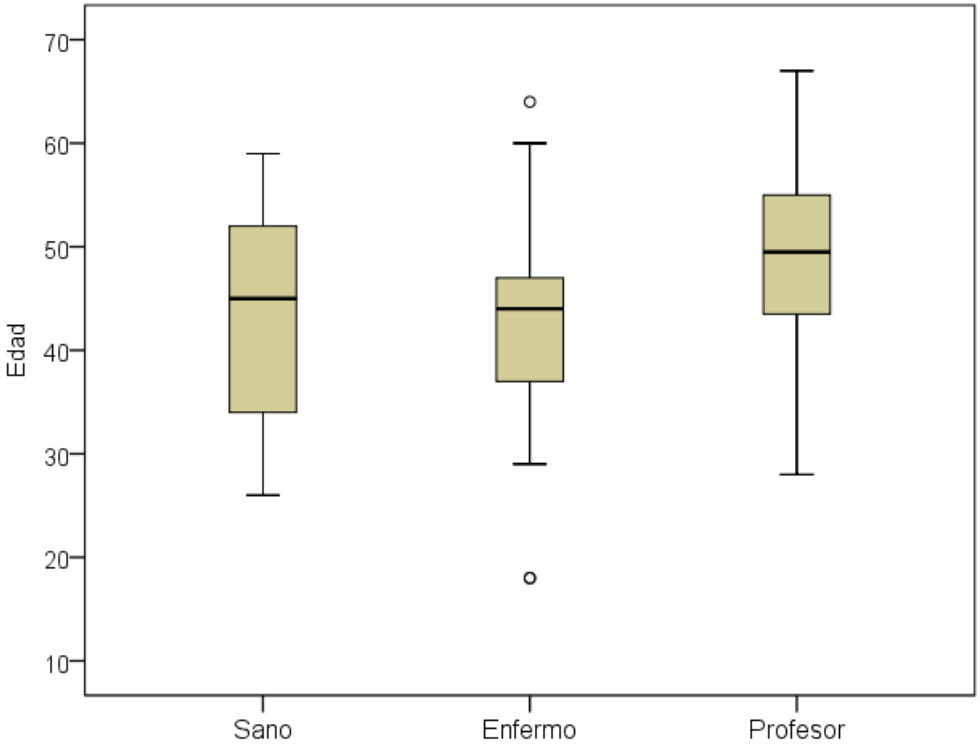

Figura 5.1. Diagrama de cajas y bigotes para edad en los distintos grupos.

Respecto al número de años fumando, los datos reflejan que el grupo de los enfermos llevaban un tiempo medio de 34,33 años (±5,13), mientras que los sanos sólo 13,25 años (±3,94). Los profesores tienen un tiempo medio como fumadores de 22,08 años (±11,90). En la Figura 5.2 se puede observar que los datos de los profesores son los más dispersos. Los enfermos presentan los valores más altos y los sanos los más bajos.

En cuanto a las puntuaciones del cuestionario VHI, se puede observar que la media más alta se presenta en el grupo de enfermos con 16,36 puntos (±10,82) frente a una puntuación muy baja del grupo de sanos con una media de 1,68 puntos (±2,70). El grupo de profesores con 3,53 puntos (±5,02) se encuentra entre los dos grupos anteriores. En el diagrama de cajas y bigotes de la Figura 5.3 se puede observar gráficamente esta situación. Además, se refleja la alta variabilidad entre los enfermos y una gran cantidad de valores atípicos en los profesores. Esta figura muestra la percepción que tienen estos tres grupos sobre su propio estado de salud vocal.

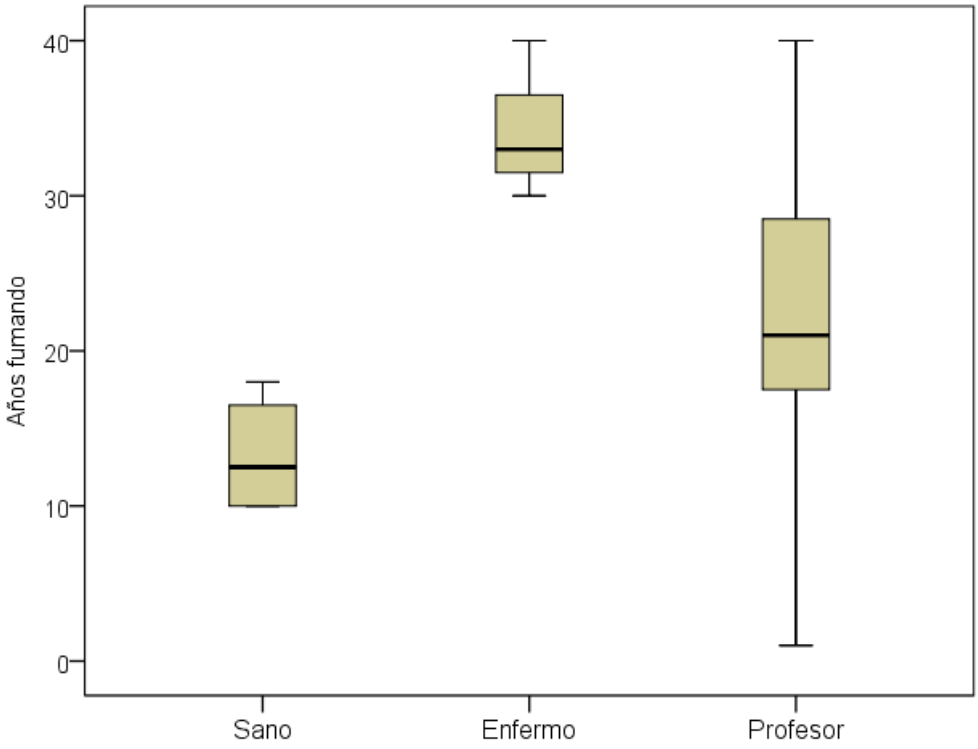

Figura 5.2. Diagrama de cajas y bigotes para el número de años fumando por grupos.

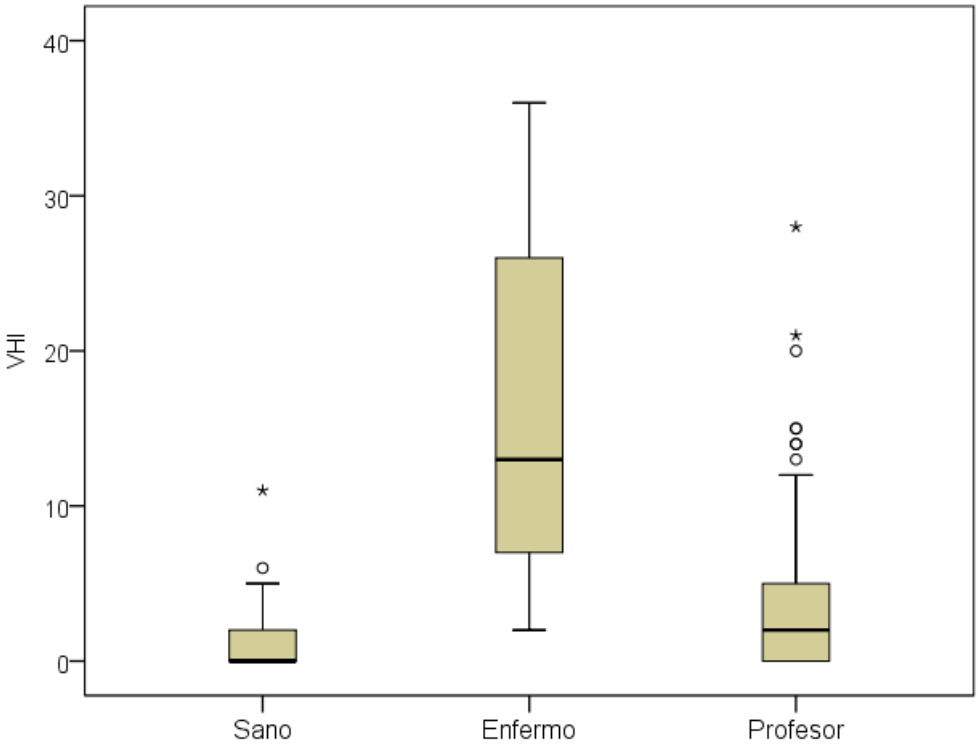

Figura 5.3. Diagrama de cajas y bigotes para las puntuaciones de VHI por grupos.

En relación con la actividad docente, el tiempo como docente de los profesores del grupo varía entre 0 y 43 años con una media de 20,77 años dedicados a la docencia (±10,93). En la Figura 5.4 se puede observar la distribución de los datos a través de un histograma. La distribución es aproximadamente simétrica.

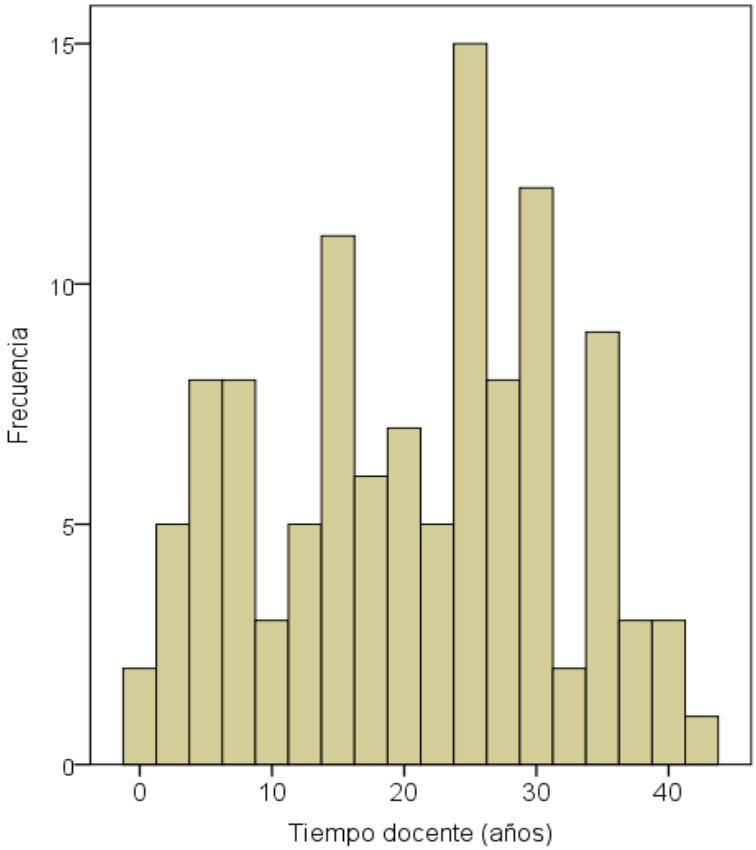

Figura 5.4. Histograma del tiempo de docencia en años del grupo de profesores.

Por último, la media de clases a la semana de este grupo es de 6,45 (±2,37). En la Figura 5.5 se representan los datos mediante un histograma. La distribución de los datos es irregular con asimetría positiva.

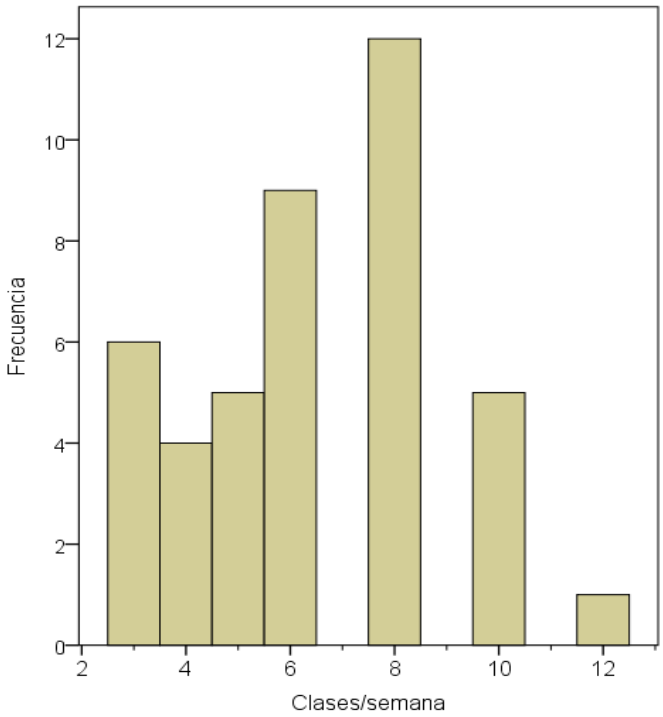

Figura 5.5. Histograma del número de horas de clases semanales del grupo de profesores.

5.2. Análisis de las características acústicas

Se han estudiado diez características que se han obtenido de las grabaciones de voz realizadas. Estas características se han dividido en dos grupos según su naturaleza: lineales (3) y no lineales (7). Entre las características lineales se encuentran la perturbación de tono local (Jitter relativo) y la perturbación de la amplitud (Shimmer local), ambas basadas en la frecuencia fundamental. La tercera característica incluida entre las lineales es la relación excitación glotal a ruido (GNE), que trata de medir la cantidad de voz debida a las cuerdas vocales frente a la debida al ruido. Esta última variable, GNE, es lineal, aunque su cálculo no se fundamente en una estimación de la frecuencia fundamental.

En la Tabla 5.3 se observan varios estadísticos descriptivos para las características acústicas lineales separados por grupo de estado.

Características		Media	Mediana	Desv.típ	Máximo	Mínimo
Jitter	Sano	0,43	0,41	0,18	0,93	0,20
	Enfermo	0,98	0,66	1,26	6,40	0,23
	Profesor	0,54	0,44	0,32	1,97	0,17
Shimmer	Sano	0,04	0,03	0,01	0,07	0,01
	Enfermo	0,06	0,05	0,03	0,13	0,02
	Profesor	0,04	0,04	0,02	0,14	0,01
GNE	Sano	0,96	0,97	0,02	0,99	0,92
	Enfermo	0,91	0,91	0,03	0,98	0,86
	Profesor	0,94	0,95	0,03	0,99	0,85

Tabla 5.3. Tabla de descriptivos para las variables lineales.

Se puede observar que tanto Shimmer como Jitter tienen valores medios (y también medianos) superiores en el caso de los enfermos. Cabe destacar también que son los enfermos los que mayor variabilidad presentan en ambos casos. En la Figura 5.6 (izquierda), se aprecia que existen valores atípicos y extremos superiores en todos los grupos para la característica Jitter relativo, siendo mayores los del grupo de enfermos y seguidos del de profesores. Los valores más bajos corresponden a los sanos, estando muy concentrados en torno a la mediana.

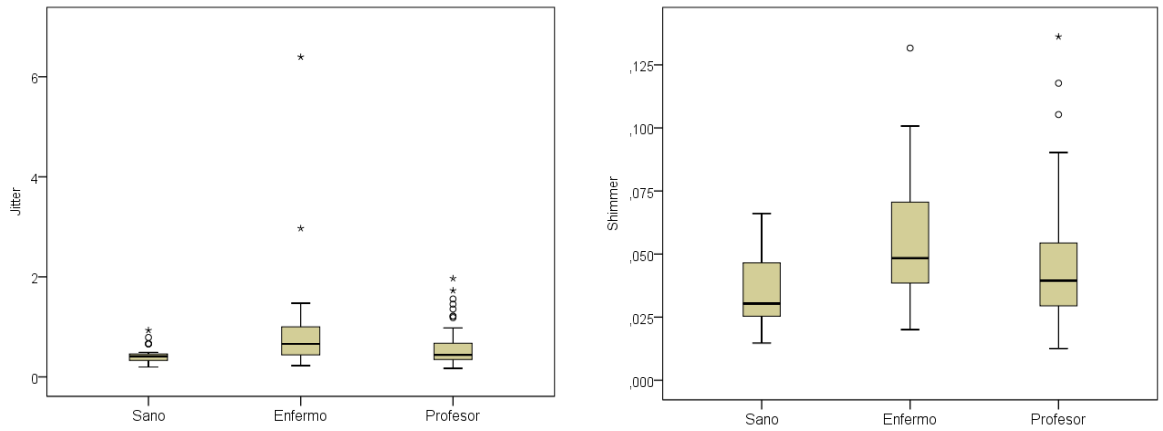

Figura 5.6. Diagrama de cajas y bigotes para Jitter relativo (izquierda) y Shimmer local (derecha) por grupos.

En el caso de Shimmer, se observan un número inferior de valores atípicos y extremos. En general, los valores más bajos corresponden a los sanos y los más altos a los enfermos, situándose los profesores entre ambos grupos. En la Figura 5.6 (derecha) puede observarse el diagrama de cajas y bigotes de la característica Shimmer.

En la Tabla 5.3 se observa que la variable GNE presenta valores más bajos y dispersos para los enfermos, mientras que los sanos se acumulan en valores altos de la distribución (asimétrica) con una dispersión menor. Como se muestra en la Figura 5.7, los profesores tienen un rango muy parecido al de los enfermos, aunque los enfermos presentan una asimetría positiva (datos más concentrados a la izquierda en valores bajos) y dos valores atípicos superiores, mientras que los profesores muestran una asimetría negativa (datos más concentrados en valores altos de la distribución) sin valores atípicos ni extremos.

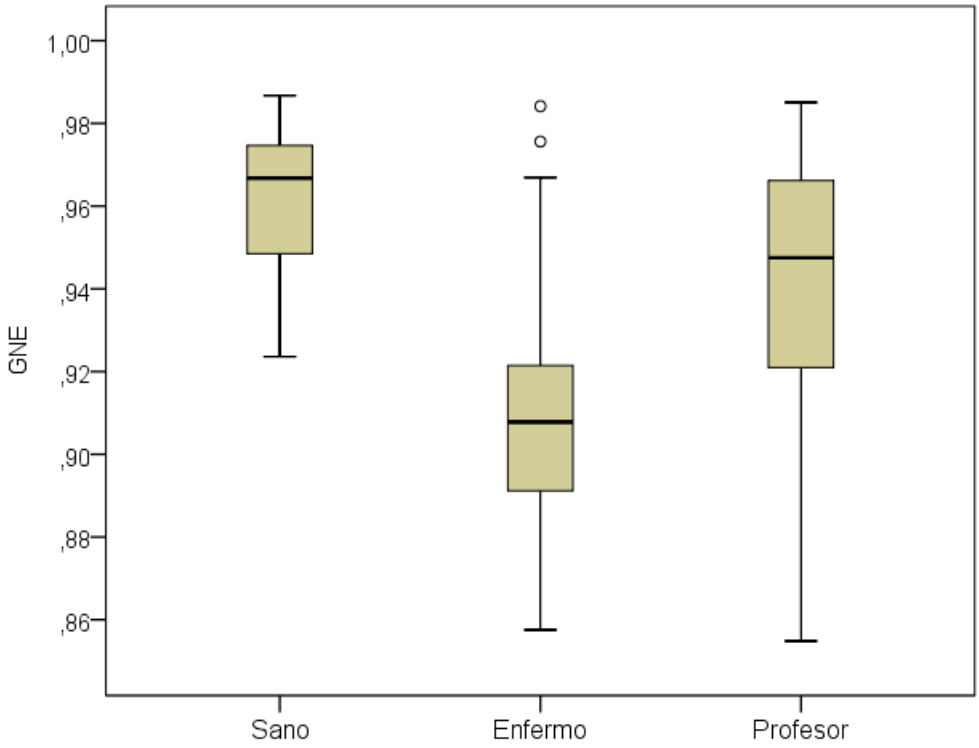

Figura 5.7. Diagrama de cajas y bigotes para GNE por grupos.

Las características no lineales tienen mucha importancia en la práctica clínica, ya que las voces disfónicas severamente patológicas son precisamente las que presentan fenómenos no lineales y aleatorios (Jiang et al., 2009). Las voces sanas están más cerca del modelo lineal de fuente filtro. Se han estudiado tres características basadas en la entropía: entropía de permutación (PermutEn), entropía difusa (FuzzyEn) y entropía muestral (SampleEn). A estas se suman otras cuatro características que son la dimensión de la correlación (D2), la entropía de densidad del período de recurrencia (RPDE), el análisis de fluctuación de tendencia (DFA) y la entropía del período de tono (PPE).

En la Tabla 5.4 se muestran algunos estadísticos descriptivos de las características acústicas no lineales. En este caso se evidencia que en el análisis de fluctuación de tendencia (DFA) son los pacientes enfermos los que tienen la media más baja (0,73).

En todas las demás variables no lineales estudiadas, los enfermos presentan los valores medios (RPDE 0,37 ±0,09; PPE 0,31 ±0,24; D2 2,51 ±0,32; PermutEn 1,36 ±0,15; FuzzyEn 0,06 ±0,04; y SampleEn 0,31 ±0,13) y medianos más altos. En general, los pacientes sanos muestran datos más concentrados en torno a la media, mientras que los enfermos son los que suelen tener mayor dispersión. El grupo de profesores también tiene valores más dispersos que los del grupo de sanos, aunque en general ligeramente por debajo de los enfermos.

Características		Media	Mediana	Desv. típ	Máximo	Mínimo
DFA	Sano	0,74	0,74	0,07	0,85	0,62
	Enfermo	0,73	0,74	0,09	0,89	0,56
	Profesor	0,77	0,77	0,08	0,96	0,52
RPDE	Sano	0,32	0,32	0,07	0,46	0,20
	Enfermo	0,37	0,35	0,09	0,55	0,23
	Profesor	0,36	0,35	0,08	0,60	0,21
PPE	Sano	0,11	0,09	0,12	0,53	0,01
	Enfermo	0,31	0,28	0,24	0,82	0,01
	Profesor	0,26	0,22	0,20	0,77	0,01
D2	Sano	2,18	2,25	0,29	2,56	1,60
	Enfermo	2,51	2,55	0,32	2,90	1,81
	Profesor	2,28	2,31	0,32	2,89	1,52
PermutEn	Sano	1,24	1,25	0,16	1,56	0,98
	Enfermo	1,36	1,37	0,15	1,66	1,03
	Profesor	1,30	1,30	0,19	1,78	0,94
FuzzyEn	Sano	0,05	0,05	0,03	0,11	0,02
	Enfermo	0,06	0,05	0,04	0,13	0,02
	Profesor	0,05	0,05	0,03	0,15	0,01
SampleEn	Sano	0,24	0,23	0,05	0,33	0,17
	Enfermo	0,31	0,30	0,13	0,72	0,15
	Profesor	0,26	0,25	0,08	0,67	0,08

Tabla 5.4. Tabla de descriptivos para las variables no lineales.

En la Figura 5.8 se aprecia la distribución de los valores de la variable DFA por grupos en un diagrama de cajas y bigotes. Esta variable es la única no lineal en la que se puede identificar que la mediana de los enfermos está ligeramente por debajo de la de los sanos. En este caso sólo se muestra un valor atípico en el grupo de profesores.

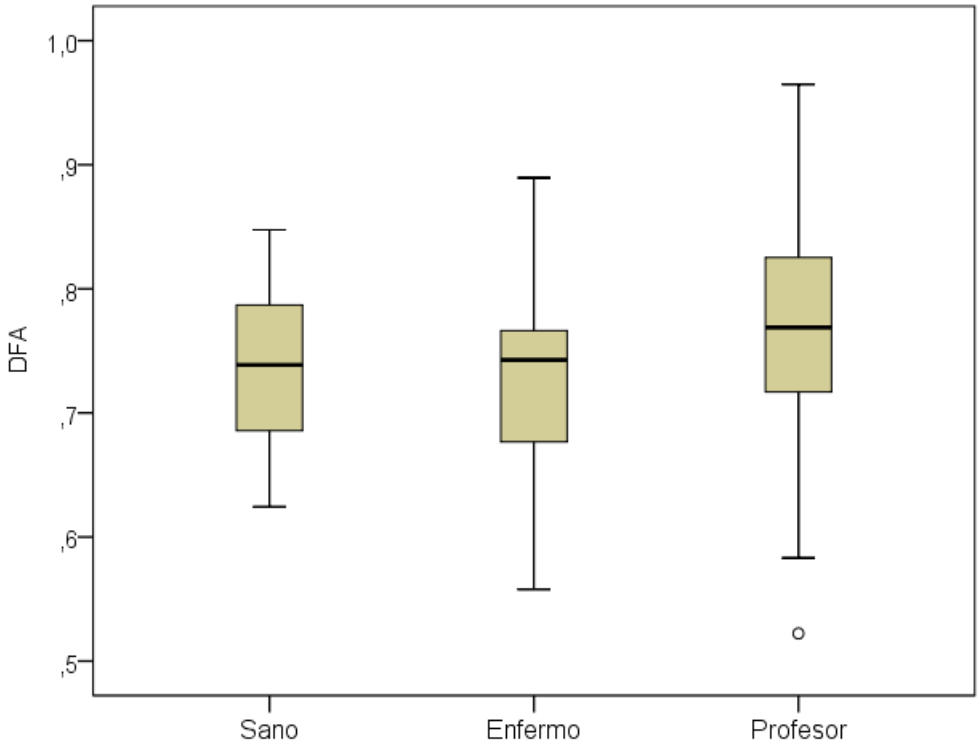

Figura 5.8. Diagrama de cajas y bigotes para DFA por grupos.

Como se observa en las Figuras 5.9, 5.10 y 5.11, tanto la mediana como los valores en general del resto de características no lineales son superiores en el grupo de enfermos. En las figuras 5.9 (derecha) (PPE) y 5.11 (izquierda) (FuzzyEn) se aprecian claras asimetrías positivas, en los tres grupos, con una cola en los valores altos de la distribución. Por otra parte, en la Figura 5.11 se puede observar una asimetría negativa en los tres grupos de la muestra, con una cola hacia valores bajos de la distribución.

La existencia de valores atípicos y extremos es diferente según los grupos y las características acústicas. Los pacientes sanos sólo muestran valores atípicos en la variable PPE (véase la Figura 5.9 derecha) y los enfermos tienen dos valores atípicos en la variable SampleEn (véase la Figura 5.11 derecha).

El grupo de profesores es el que más valores atípicos presenta. En las figuras 5.9 (izquierda) (RPDE) y 5.11 (izquierda) (FuzzyEn) se pueden observar valores atípicos superiores para este grupo, mientras que en la Figura 5.11 (derecha) (SampleEn) pueden observarse varios valores atípicos superiores, así como un valor atípico inferior.

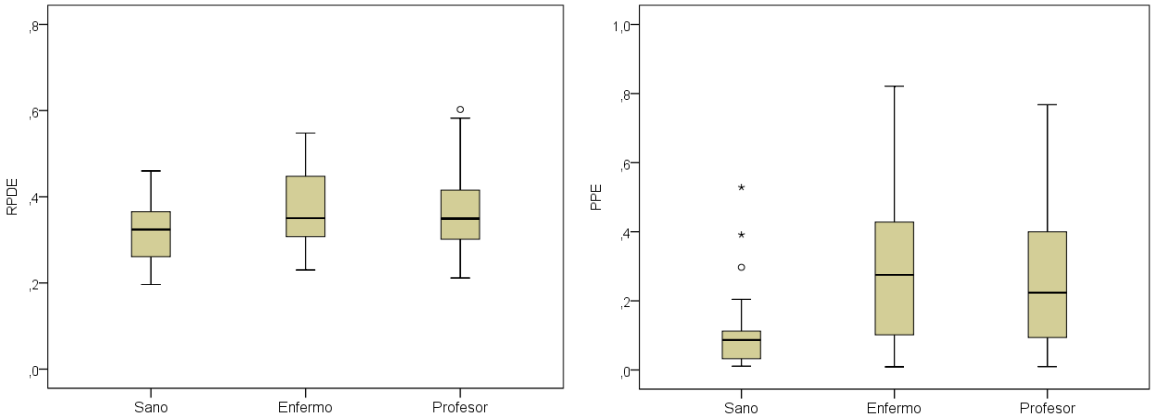

Figura 5.9. Diagrama de cajas y bigotes para RPDE (izquierda) y PPE (derecha) por grupos.

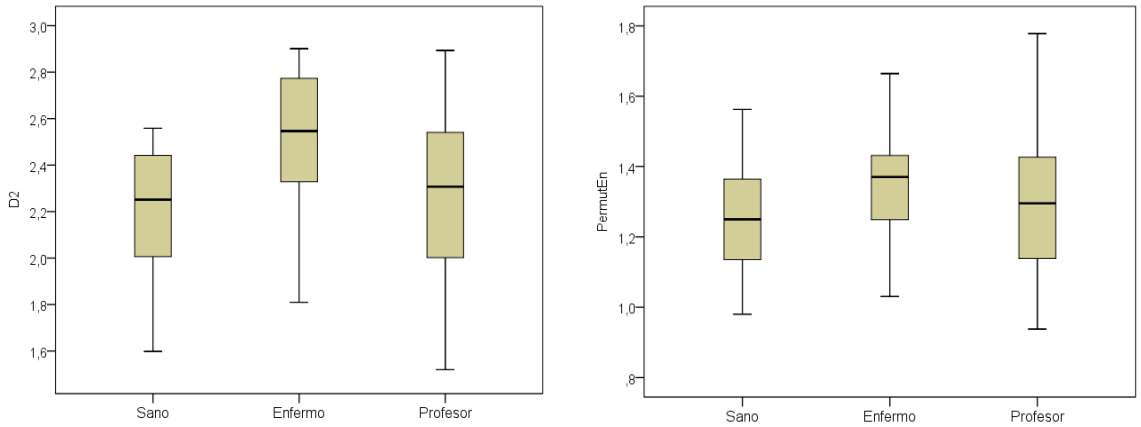

Figura 5. 10 Diagrama de cajas y bigotes para D2 (izquierda) y PermutEn (derecha) por grupos.

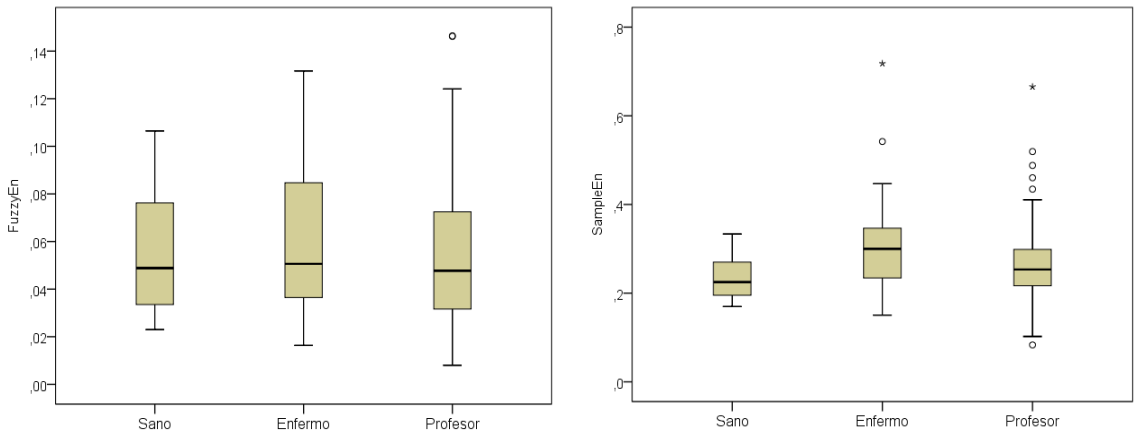

Figura 5.11. Diagrama de cajas y bigotes para FuzzyEn (izquierda) y SampleEn (derecha) por grupos.

5.3. Aplicación de modelos predictivos

En esta sección se estudia la capacidad predictiva de las variables acústicas consideradas tras un proceso de selección de variables y clasificación. Generalmente, para ajustar modelos de clasificación y predicción se busca la simplicidad sin perder precisión. Esto permite obtener una explicación más intuitiva de los modelos utilizados, consiguiendo un nivel de acierto suficientemente bueno en las predicciones. La utilización de un número elevado de características para este tipo de análisis puede provocar problemas de multicolinealidad o sobreajuste en los modelos de clasificación, por lo que el número de características debe permanecer lo más bajo posible siempre que se mantenga un nivel de explicación de la variable dependiente suficientemente alto con las variables explicativas que se mantengan en el modelo. Esto se conoce como el principio de parsimonia (Tsanas & Gómez-Vilda, 2013). Para conseguir este equilibrio, se ha utilizado el método discriminante lineal de Fisher (Al-Nasheri et al., 2017).

El mejor subconjunto de variables, partiendo de las diez iniciales y siguiendo las indicaciones anteriores y explicadas en el Capítulo 4, está formado por: SampleEn, D2, PPE, GNE y sexo. Incluyendo únicamente estas variables se alcanzó una proporción de acierto de 0,92 (intervalo de confianza al 95%, 0,86-0,98) con el clasificador Random Forest. Tanto la sensibilidad como la especificidad también fueron 0,92. Aumentar el número de características acústicas en este caso no aumenta la proporción de acierto. La Tabla 5.5 muestra las métricas del rendimiento de los once modelos de clasificación utilizados, incluyendo el seleccionado (Random Forest).

Clasificador	Prop. acierto	AUC	Sensibilidad	Especificidad
KNN	0,76	0,80	0,56	0,96
RF	0,92	0,94	0,92	0,92
VGLAC	0,76	0,87	0,72	0,80
ERGLM	0,82	0,92	0,76	0,88
ANN	0,74	0,86	0,72	0,76
SVMP	0,84	0,90	0,80	0,88
BCRT	0,90	0,92	0,84	0,96
BLR	0,86	0,90	0,90	0,92
BART	0,84	0,86	0,76	0,92
CRF	0,88	0,94	0,80	0,96
GLME	0,89	0,95	0,89	0,88

Tabla 5.5. Métricas estadísticas del rendimiento del clasificador con los predictores seleccionados.

Una vez seleccionadas las variables y estudiados los resultados de clasificación conseguidos con estas variables y los once modelos, resulta de especial interés comprobar si existen diferencias estadísticamente significativas entre los enfermos y los sanos. Como se observa en los diagramas de cajas y bigotes de la Figura 5.12, los enfermos tienen una mediana y valores, en general, superiores a los de los sanos en tres de las cuatro variables acústicas seleccionadas. En la variable GNE los sanos obtienen los valores mayores.

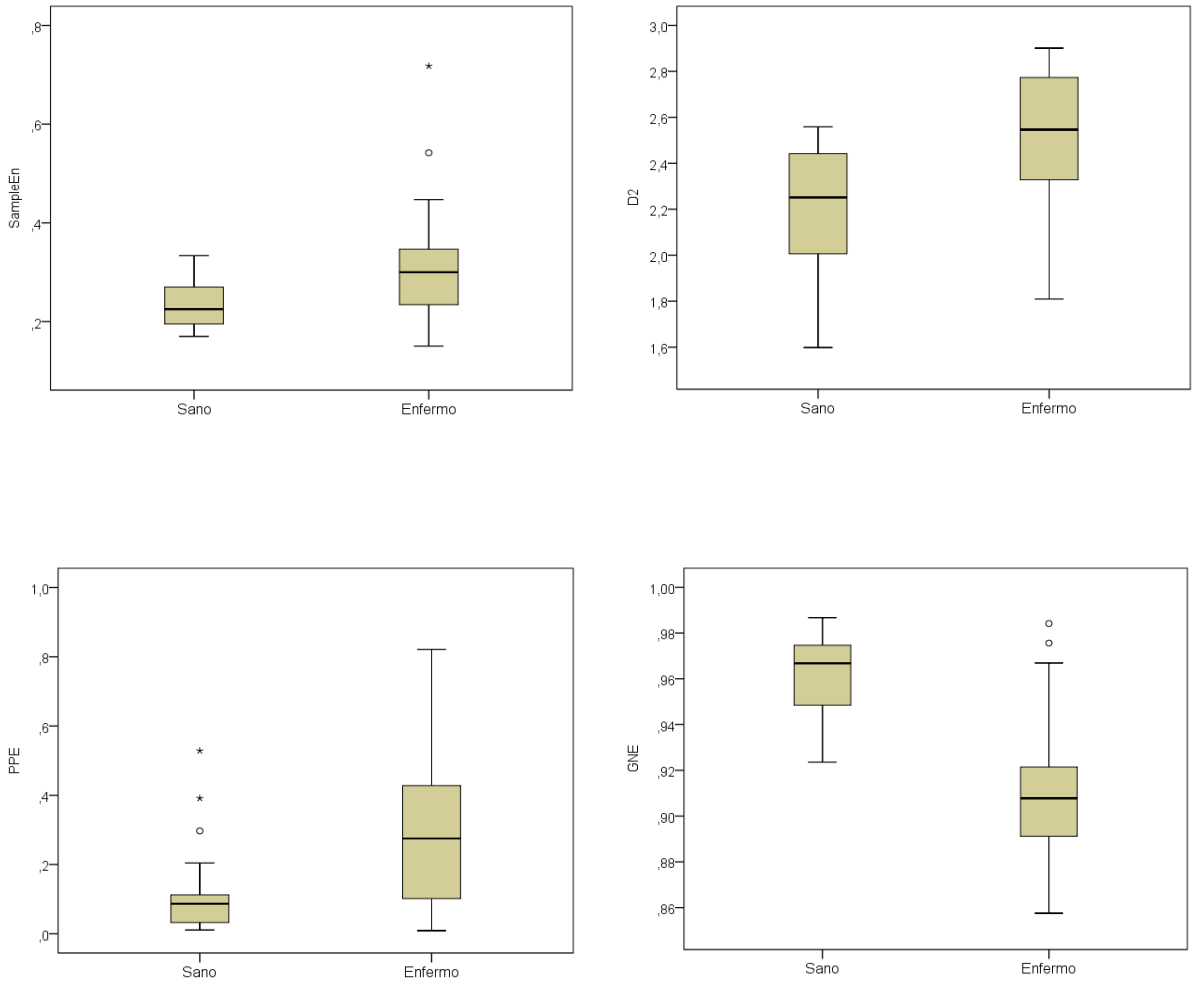

Figura 5.12. Diagrama de cajas y bigotes de las variables SampleEn (arriba izquierda), D2 (arriba derecha), PPE (abajo izquierda), GNE (abajo derecha).

Las características acústicas de los pacientes sanos frente a los enfermos se han analizado mediante la realización de contrastes de hipótesis de comparación de medias. En todos los casos se encontraron diferencias estadísticamente significativas (p-valores<0,05) entre los grupos de sanos y enfermos (véase la Tabla 5.6), destacando así la capacidad discriminadora de las variables elegidas y apoyando esta selección como adecuada para diferenciar entre ambos grupos.

Características		Media	Mediana	Desv. típ	Mínimo	Máximo	p-valor
SampleEn	Enfermos	0,31	0,30	0,13	0,15	0,71	0,018
	Sanos	0,24	0,23	0,05	0,17	0,33	
D2	Enfermos	2,51	2,55	0,32	1,81	2,90	<0,001
	Sanos	2,18	2,25	0,29	1,60	2,55	
PPE	Enfermos	0,31	0,78	0,24	0,01	0,82	0,002
	Sanos	0,11	0,09	0,12	0,01	0,53	
GNE	Enfermos	0,91	0,91	0,03	0,86	0,98	<0,001
	Sanos	0,96	0,97	0,17	0,92	0,99	

Tabla 5.6. Estadísticos descriptivos y p-valores para las pruebas de comparación de medias de las características acústicas seleccionadas.

5.4. Asignación de profesores a niveles de prevención

En esta sección se presentan, en primer lugar, algunos resultados comparativos. Posteriormente, se analizan los resultados del proceso de aprendizaje automático, diseñado para asignar al personal docente a los diferentes niveles de prevención de patologías de la voz.

La autoevaluación de la salud vocal a partir del cuestionario realizado en los distintos grupos (VHI) aportó diferentes valores medios para sanos y enfermos. El grupo que mayor media obtuvo fue el de enfermos, con 16,36 puntos (± 10,82), mientras que el grupo de sanos tuvo una media de 1,68 puntos (±2,7). Se detectaron diferencias estadísticamente significativas entre el grupo de enfermos y el de sanos (p-valor < 0,001). Si incluimos en el análisis el grupo de profesores, sobre el que queremos hacer predicciones, se observan diferencias estadísticamente significativas en las puntuaciones medias del cuestionario entre el personal docente y el grupo de enfermos (p-valor < 0,001), pero no entre las del personal docente y las del grupo de sanos (p-valor = 0,172). Por lo tanto, y en términos generales, las puntuaciones obtenidas en el cuestionario por los profesores están más próximas a las mostradas en el grupo de sanos que a las del grupo de enfermos. La Figura 5.13 muestra los diagramas de caja y bigotes correspondientes a los tres grupos analizados en la muestra. Los valores atípicos en el caso del personal docente requieren atención especial, ya que indican una autopercepción negativa de las voces de algunos profesores concretos.

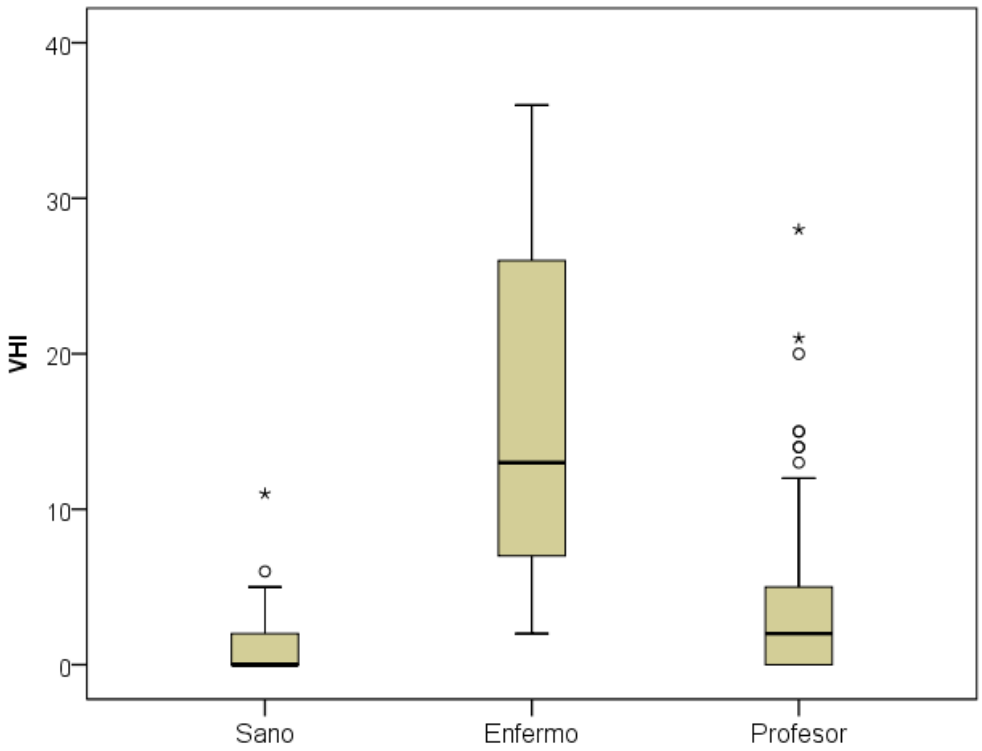

Figura 5.13. Diagrama de cajas y bigotes para VHI por grupos.

Mediante la aplicación de técnicas de Machine Learning, como es el clasificador Random Forest, se ha conseguido un modelo que permite discriminar con alta presión entre individuos sanos e individuos con nódulos en las cuerdas vocales (u otras patologías vocales con similares características). Basándose en los parámetros obtenidos mediante el aprendizaje con sujetos sanos y enfermos, se aplicaron sobre las variables seleccionadas (SampleEn, D2, PPE, GNE y sexo) a los profesores para estimar la probabilidad de que tengan nódulos en las cuerdas vocales u otras patologías vocales con características similares. Con estas probabilidades se realizó la asignación a los tres grupos de riesgo como se expone el Capítulo 4. El resultado obtenido mostró que el 60,8% (73 de 120) de los profesores debían ser asignados al grupo de riesgo bajo, el 29,2% (35 de 120) al grupo de riesgo medio, y el 10% (12 de 120) al grupo de riesgo alto.

Mediante el análisis del cuestionario realizado por todos los profesores, se observó que 17 de los 120 profesores, reconocieron la existencia de problemas en su voz (véase la Tabla 5.7). Se observa una asociación estadísticamente significativa entre la variable de reconocimiento de problemas vocales con el grupo de riesgo asignado (p-valor < 0,001), demostrando de una relación de dependencia (V de Cramér = 0,487).

Problemas de voz x grupo de riesgo			Grupo de Riesgo			
			Bajo	Medio	Alto	Total
Problemas de voz	No	Recuento	72	25	6	103
		% por problemas	69,90	24,27	5,83	100
	Si	Recuento	1	10	6	17
		% por problemas	5,88	58,82	35,29	100
	Total	Recuento	73	35	12	120
		% por problemas	60,83	29,17	10,00	100

Tabla 5.7. Tabla de contingencia para la variable problemas de voz y grupo de riesgo asignado.

Casi el 70% de los profesores que afirmaron no tener problemas vocales fueron clasificados en el grupo de bajo riesgo, seguidos por el 24,27% que fueron clasificados como de riesgo medio.

Con estos resultados se pudo aplicar un plan de prevención. De este modo, los profesores pertenecientes al grupo de riesgo bajo se incorporaron a un programa de salud de atención primaria, mientras que los asignados al grupo de riesgo medio siguieron programas de atención secundaria. Los seis profesores (3 hombres y 3 mujeres) que no reconocieron problemas en su voz, pero fueron asignados por el modelo al grupo de alto riesgo ya que no presentaron grabaciones de voz sugestivas de patología vocal. A pesar de esto, la presencia de diversos factores de riesgo (elevado número de años en el desempeño de la labor docente, fumadores, puntuaciones altas en cuestionario VHI o episodios previos de trastornos vocales) en dos de ellos, supuso la necesidad de valoración por un especialista en ORL, y el seguimiento por parte del Servicio de Prevención de la Universidad de Extremadura.

La mayoría de los profesores que reconocieron tener problemas de voz se clasificaron en el grupo de riesgo medio (58,82%), seguidos de un 35,29% que fueron clasificados en el grupo de riesgo alto. Sólo uno de los participantes que reconoció tener problemas vocales fue asignado al grupo de riesgo bajo por el clasificador (5,88%). Al revisar el cuestionario respondido por este sujeto, se comprobó que padecía una disfonía por rinofaringitis crónica, por lo que debería haber sido incluido en el grupo de riesgo alto. Debido a la existencia de una patología crónica en estructuras que conforman el aparato fonador, se le incluyó en un programa de prevención terciaria.

Respecto al grupo de riesgo medio (7 hombres y 3 mujeres), todos indicaron que tenían una afectación de la voz relativamente leve, como disfonía o irritación faringo-laríngea, con varios episodios durante los últimos años.

Este hecho, conllevó la aplicación de un programa de atención secundaria para estos sujetos. Los participantes clasificados en el grupo de riesgo alto (5 hombres y una mujer) presentaron diferentes problemas vocales. De los seis, uno de ellos afirmó tener una irritación faringo-laríngea de carácter agudo, y no haber presentado episodios previos similares, por lo que finalmente se le incluyó en un programa de atención primaria y seguimiento. Tres sujetos sufrían disfonía crónica con episodios repetidos en los últimos años, y se incorporaron a un programa de atención terciaria. Los dos participantes restantes padecían ciertamente nódulos vocales, por lo que se les asignó también a programas de prevención terciaria.

Una vez que se establecieron los niveles de prevención, se estudió la posible asociación estadística entre el estado de voz reconocido y el nivel de prevención asignado (véase la Tabla 5.8). Nuevamente, se encontró una asociación estadísticamente significativa entre las dos variables ($p<0,001$), con una clara relación de dependencia (V de Cramer = 0,630).

Problemas de voz x Niveles de prevención			**Niveles de prevención**			
			Primaria	**Secundaria**	**Terciaria**	**Total**
Problemas de voz	No	Recuentos	76	27	0	103
		% por problemas	73,78	26,21	0	100
	Si	Recuentos	1	11	5	17
		% por problemas	5,88	64,71	29,41	100
	Total	Recuentos	77	38	5	120
		% por problemas	64,16	31,67	4,17	100

Tabla 5.8. Tabla de contingencia para la variable problemas de voz y los niveles de prevención asignados.

Como con cualquier método de diagnóstico, esta asignación automática a grupos de riesgo puede generar clasificaciones erróneas. Por este motivo la valoración clínica basada en una anamnesis rigurosa, junto con una exploración a través de pruebas de imagen como son el fibrolaringoscopio y estroboscopio, son fundamentales en el diagnóstico de la patología vocal. Sin embargo, poder cribar trastornos vocales a través de técnicas no invasivas y de bajo coste en servicios de prevención supone un beneficio para los pacientes y para la administración.

En el caso de los primeros, porque podrán llevarse a cabo actuaciones tempranas en el caso de detectarse problemas en estadios iniciales. En el caso de la administración, porque podrán evitarse o cuando menos reducirse los costes de bajas laborales por trastornos vocales evolucionados y/o no tratados.

A partir de estas asignaciones, se pueden desglosar resultados por niveles de prevención. La Tabla 5.9 presenta la relación entre los niveles de prevención asignados y el sexo. Únicamente el 4,17% del total del profesorado fueron asignados al nivel de prevención terciario. En Prevención primaria se asignó un porcentaje menor de hombres que de mujeres (61,33% vs 68,89%) mientras que en prevención secundaria el porcentaje fue mayor en hombres que mujeres (34,67% vs 26,67%). En cualquier caso, no se detectaron evidencias suficientes para determinar que las diferencias entre sexos fueran estadísticamente significativas (p-valor=0,660).

Sexo x Niveles de prevención			**Niveles de prevención**			
			Primaria	**Secundaria**	**Terciaria**	**Total**
Sexo	Hombre	Recuentos	46	26	3	75
		% por sexo	61,33	34,67	4,00	100
	Mujer	Recuentos	31	12	2	45
		% por sexo	68,89	26,67	4,44	100
	Total	Recuentos	77	38	5	120
		% por sexo	64,17	31,16	4,17	100

Tabla 5.9. Tabla de contingencia de los niveles de prevención en los distintos sexos

En la Tabla 5.10 se presenta relación entre la asignación a niveles de prevención y la condición de fumador. En prevención primaria destacan los no fumadores (64,81% vs 58,33%), mientras que en prevención secundaria son mayoría los fumadores (30,56% vs 41,67%). El porcentaje de profesores asignados a terciaria fue de un 4,63% y todos ellos no fumadores. En el nivel de prevención terciaria no hubo fumadores.

El tiempo como fumador puede ser un factor favorecedor para padecer patologías relacionadas con la garganta. En este caso, se observa que el número medio de años como fumador entre los profesores asignados a prevención primaria es de 16,86±10,89, mientras que entre los asignados a prevención secundaria la media sube a 29,4 años (±9,84).

Fumador x Niveles de prevención			Niveles de prevención			
			Primaria	Secundaria	Terciaria	Total
Fumador	No	Recuentos	70	33	5	108
		% por sexo	64,81	30,56	4,63	100
	Sí	Recuentos	7	5	0	12
		% por sexo	58,33	41,67	0	100
	Total	Recuentos	77	38	5	120
		% por sexo	64,17	31,67	4,17	100

Tabla 5.10. Tabla de contingencia de los niveles de prevención según condición de fumador

En la Figura 5.14 se presenta un diagrama de cajas y bigotes para la puntuación VHI-10 según nivel de prevención asignado.

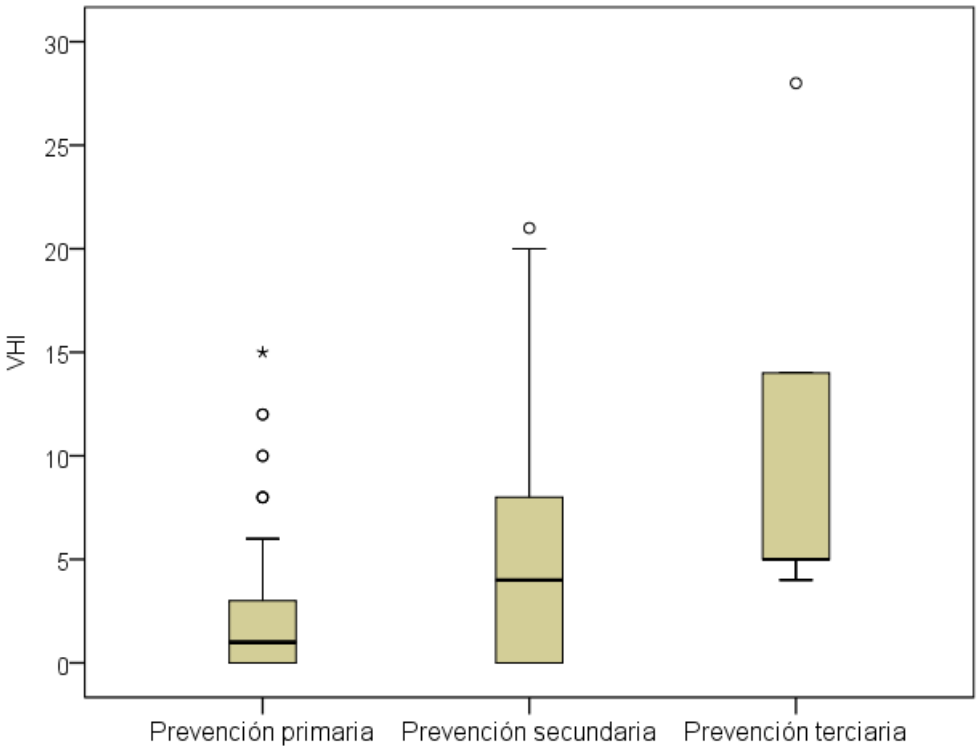

Figura 5.14. Diagrama de cajas y bigotes para VHI por nivel de prevención asignado.

Se destaca que la mediana del índice VHI-10 que obtuvieron los profesores asignados al nivel de prevención primaria fue la más baja, mientras que la de los asignados al nivel de prevención terciaria fue la más alta. De manera general, los valores de VHI de los profesores asignados a prevención primaria son más bajos y concentrados que el resto, aunque muestran un valor extremo y tres valores atípicos superiores. En prevención secundaria se aprecian valores muy dispersos, con gran cantidad de valores iguales a cero como ocurre en prevención primaria. También incluye valores de VHI-10 altos, como algunos de los que se obtienen en prevención terciaria. Tanto en prevención secundaria como terciaria se observa un valor atípico superior. Las distribuciones de VHI-10 en los tres niveles son asimétricas positivas.

Capítulo 6. Discusión

Con el presente trabajo de tesis doctoral se ha demostrado que es posible desarrollar y aplicar un programa de prevención de enfermedades de la voz a partir de características acústicas extraídas automáticamente de grabaciones de voz. A partir de la clasificación automática de profesores en diferentes grupos de riesgo de padecer nódulos vocales u otras patologías vocales similares, se han llevado a cabo acciones de promoción, prevención y rehabilitación de la salud vocal con el fin de retrasar, mejorar o disminuir los signos y/o síntomas que conllevan el desarrollo de trastornos vocales como los descritos.

Todas las patologías vocales no pueden ser tratadas y/o corregidas a través de actuaciones conservadoras. Sin embargo, las acciones encaminadas a la educación, desarrollo y potenciación de hábitos y técnicas vocales saludables contribuyen a mejorar la calidad vocal (y de vida) de los docentes afectados. Este planteamiento coincide con el de otros expertos que describen recomendaciones y normativas que regulan las condiciones de los ambientes de trabajo de los docentes, así como las actuaciones a realizar sobre ellos una vez diagnosticado problema vocal (Santana et al., 2012; De Oliveira Bastos, & Hermes, 2018; Kyriakou et al., 2018).

La OIT considera a los docentes como el primer grupo de riesgo de padecer enfermedades profesionales de la voz (Behlau, et al., 2012; Farías, 2013; Niebudek-Bogusz & Śliwińska-Kowalska, 2013). En la legislación española el concepto de enfermedad profesional viene definido en el artículo 157 del texto refundido de la Ley General de la Seguridad Social aprobado por el Real Decreto Legislativo 8/2015 de 30 de octubre. Sin embargo, no es hasta el año 2006 cuando las enfermedades de la voz, en concreto los nódulos en las cuerdas vocales, se reconocen en el cuadro de enfermedades profesionales con el código 2L0101 (RD 1299/2006 de 10 de noviembre) dentro del grupo “Enfermedades producidas por agentes físicos”, precisando como actividades capaces de producirlos “aquellas actividades en las que se precise un uso mantenido y continuo de la voz, como son profesores, cantantes, actores, teleoperadores, locutores”. Este aspecto, y el hecho de que los nódulos vocales constituyen una de las patologías vocales más comunes entre el personal docente, han sido las principales razones para considerar esta patología como eje principal en el desarrollo del programa propuesto.

A lo largo de los años, en la literatura médica, diferentes autores se han interesado por la presencia de nódulos vocales entre los docentes. Urrutikoetxea et al. (1995) observaron que, de una muestra de 1046 docentes, 218 sujetos presentaban lesiones vocales orgánicas. De los que presentaban este tipo de lesiones, 94 padecían nódulos vocales (43,12%), 39 edema de Reinke, 24 hipertrofia de bandas vestibulares, 19 pólipos y el resto otras lesiones menos frecuentes. Una proporción mayor fue la encontrada por Tavares y Martins (2007), que observaron que el 63,16% de las lesiones orgánicas encontradas entre un grupo de 80 docentes analizados eran nódulos vocales.

Arrizabalaga y García (2017) observaron que en una muestra de 70 individuos que acudieron a consulta médica en servicios de atención primaria en la comunidad de Navarra refiriendo padecer una disfonía, el 81,7% fueron docentes y, de esos, finalmente el 31% fue diagnosticado con nódulos vocales. Preciado et al. (2005), en un estudio realizado a 492 docentes en la comunidad de La Rioja, para analizar los factores de riesgo que influyen en la aparición de trastornos vocales, observaron que el 13,8% de las lesiones diagnosticadas eran nódulos vocales. En el estudio presentado, mediante un modelo de clasificación automática y sin la necesidad de realizar un diagnóstico clínico de patologías vocales, el 58,8% de los profesores fue clasificado en el grupo de riesgo medio y el 35,3% en el grupo de riesgo alto de padecer nódulos vocales u otras patologías compatibles con los mismos.

La edad puede estar relacionada con el momento de aparición de este tipo de enfermedades de la voz. Según los estudios realizados por diferentes autores, como Elhendi (2005), Elhendi et al. (2012b), Cervera et al. (1996), o Pérez et al., 2003), se presenta una mayor incidencia de este tipo de patología vocal en la cuarta década de la vida, seguida por la tercera. En este estudio, se puede observar que la edad media del grupo de profesores es de 48,91 años (± 8,51), y entre ellos, los asignados a los niveles de prevención secundario y terciario son de 48,76 años (±8,99) y 49,6 años (±5,68), respectivamente. Autores como Herrington-Hall et al. (1988) y Le Huche y Allalí (1994) coinciden en afirmar que la mayor incidencia de estos trastornos se encuentra en la tercera década de la vida, mientras que en el experimento presentado se encuentran en la cuarta década.

En nuestro caso, la muestra de enfermos de nódulos se seleccionó según la llegada a consulta de disfonía en el periodo que se realizó el experimento. La muestra de sanos se eligió para que fuera equilibrada respecto al sexo con respecto a la de enfermos, por lo que el porcentaje de hombres es del 16% y el de mujeres es de un 84% en ambos casos. Este porcentaje más elevado de mujeres en el grupo de enfermos (pacientes con diagnóstico firme de nódulos vocales) puede venir determinado por la mayor prevalencia de nódulos vocales en mujeres, y coincide con los datos encontrados en la literatura (Preciado et al., 2005; Van Houtte, 2010; Elhendi et al., 2012c; Arrizabalaga & García, 2017). Algunos autores explican este hecho basándose en las diferencias anatómicas que existen entre las laringes de hombres y mujeres. Dejonckere y Kob en 2009 establecieron que los bordes libres de las laringes femeninas eran más curvados que los de los hombres, hecho que favorecía un microtraumatismo repetido durante la fonación. Para ello, realizaron simulaciones en tres dimensiones del ciclo vocal, estudiando las características anatómicas de las laringes femeninas en cada momento del ciclo. Otros autores, justifican la presencia de porcentajes elevados de lesiones vocales por nódulos en mujeres, a que las mujeres utilizan con más frecuencia un patrón respiratorio de predominio torácico superior, fijando y tensionando la musculatura del cuello y de los hombros, siendo esta una actitud inadecuada para una adecuada fonación (véase, por ejemplo, Bustos, 1991).

Para autores como Menezes et al. (2011), el hecho de que la disfonía se diagnostique con más frecuencia en el sexo femenino está relacionado, por una parte, con una mayor preocupación de la mujer por su salud, que hace que consulte con mayor frecuencia y rapidez a un especialista y, por otra, con un mayor sobreesfuerzo vocal realizado con respecto a los hombres. Se puede observar que, en este experimento, entre los profesores, el porcentaje de hombres es de un 62,5% (75) frente a un 37,5% (45) de mujeres. Se debe tener en cuenta que no se trata de un muestreo estadístico, sino de los profesores que voluntariamente se realizaron el examen de salud periódico de la Universidad de Extremadura. Del total de hombres de la muestra, 26 fueron asignados al nivel de prevención secundario (34,67%) y 3 al terciario (4%), mientras que la mayoría se asignaron al nivel de prevención primaria (61,33%). En el caso de las mujeres, 12 fueron asignadas al nivel de prevención secundaria (26,67%) y 2 al nivel terciario (4,44%), mientras que, también en este caso, la mayoría se asignó al nivel primario (68,89%). Estos datos muestran un porcentaje mayor de hombres asignados al nivel de prevención secundaria, mientras que el porcentaje de mujeres asignadas al nivel terciario es ligeramente superior, aunque las proporciones se encuentran bastante equilibradas.

La forma en que nos relacionamos y los estilos de vida pueden afectar en mayor o menor medida a la salud vocal. Para Castejón (2013) en este proceso participan tres componentes: el profesor y su circunstancia individual (hábitos, personalidad, estilos de vida), el contexto en el que desarrolla su trabajo (ruido, temperatura, acústica) y la interacción del profesor con el contexto de trabajo (regulación del tono, volumen, rapidez de la voz en distintos espacios). La existencia o no de hábito tabáquico y su temporalidad entran dentro del estilo de vida. Si bien es cierto que la literatura no considera el tabaco como un factor etiológico en el desarrollo de nódulos vocales (sí lo es en otras patologías como el edema de Reinke o el carcinoma de laringe), se contempla como factor favorecedor de su aparición (Le Huche, 1994; Urrutikoetxea, 1995; Elhendi, 2005). En los grupos de sanos y de enfermos el porcentaje de individuos que afirman ser fumadores es del 16%, con una media de años fumando de 34,33 (±5,13) en el caso del grupo de enfermos y de 13,25 (±3,94) en el grupo de los sanos; mientras que en el de profesores este porcentaje es del 10% y la media de años fumando es de 22,08 (±11,90), cantidad que se encuentra entre las medias de sanos y enfermos, pero que muestra una mayor dispersión. Además, en el caso de los sanos, el número de años fumando es casi la mitad de lo que muestran los datos del grupo de enfermos. Cabe destacar que en el grupo de profesores la variación del número de años fumando se encuentra entre 1 y 40, mientras que en los grupos de sanos se encuentran en 10-18 y 30-40, respectivamente. La presencia de valores tan elevados en el grupo de enfermos puede sugerir que el hábito tabáquico puede ser un agente favorecedor en la aparición de los trastornos vocales estudiados. Además, el mayor porcentaje de exfumadores se da entre los enfermos, con un 52%, seguido del grupo de profesores con el 33,6% y del de sanos con un 26,3%. Esto puede ser debido a las recomendaciones de hábitos saludables aportadas por los equipos sanitarios una vez detectada alguna patología de la voz.

En cuanto a los niveles de prevención, del 10% de profesores que afirman fumar (12 en total), 5 (41,67%) fueron asignados al nivel de prevención secundaria y ninguno al terciario. Este alto porcentaje en el nivel de prevención secundaria es indicativo también de que el hábito tabáquico puede ser un agente favorecedor.

Por otra parte, el abordaje del hábito tabáquico desde el punto de vista de la temporalidad en años de consumo no queda recogido suficientemente en análisis epidemiológicos en la literatura consultada. A pesar de que no se encontraron diferencias estadísticamente significativas entre el número medio de años fumando en los niveles de prevención primaria y secundaria (no se asignó a ningún profesor fumador al nivel de prevención terciaria), se puede apreciar que, en el nivel de prevención primaria, la media de tiempo de fumador bajó de 22,08 (±11,90) en todos los profesores a 16,86 años (±10,89), mientras que en el nivel de prevención secundaria subió a 29,40 años (±9,84). La falta de significación estadística se debe al reducido tamaño muestral. Otros autores contemplan el hábito tabáquico en relación con el consumo de número de cigarrillos/día (Fernández & Preciado, 2003) o simplemente basándose en la condición de fumador o no fumador y no considerando el tiempo que llevan fumando. En este sentido, Elhendi et al. (2012) muestran que el 27,6% de los sujetos encuestados que presentaban disfonía eran fumadores, frente al 72,3% que no lo eran, pero no encuentran relación significativa entre el consumo de tabaco y la aparición de la afección. Esto coincide con lo aportado por López y Cervilla (2016) y Gañet et al. (2007). La condición de exfumador y el tiempo sin fumar de los mismos es un dato que no viene reflejado en los estudios consultados. En este trabajo presentado no se aprecian diferencias estadísticamente significativas entre el nivel de prevención asignado y la condición exfumador. Sin embargo, se ha obtenido que el porcentaje de exfumadores es mayor en el nivel de prevención secundaria (40,5%) que en el de prevención primaria (32,4%).

La carga vocal es un concepto que está marcado por la demanda a que se somete el sistema vocal. Gorospe et al. (2004) definieron una serie de factores que intervienen como factores directos en dicha demanda: tiempo de fonación durante la jornada de trabajo, cuyo índice es el porcentaje de tiempo medio de fonación por hora de trabajo; la intensidad de habla; el tono medio utilizado al hablar; el nivel de ruido ambiental y el tiempo de descanso vocal. Sin embargo, existen otros aspectos vinculados al ámbito personal del individuo que inciden de forma indirecta en la fonación, y que pueden aumentar el riesgo de disfonía. Entre ellos destacan circunstancias relacionadas con el espectro emocional (estrés, personalidad, ansiedad…), condiciones de salud como enfermedad aguda o crónica o debilidad generalizada, o hábitos nutricionales, de consumo de sustancias tóxicas (alcohol, tabaco, drogas…) y de vida (sedentarismo, sueño insuficiente…) (Gotaas & Starr, 1993; Le Huche & Allalí, 2003). La carga vocal no es igual en los diferentes niveles educativos. Sodersten et al. (2002) afirman que en el caso de los maestros de educación infantil la carga está determinada por 5 horas de clase diarias con un promedio de tiempo del 20% de fonación por hora de trabajo y una intensidad media >80dB. Por otro lado, Buekers et al. (1995) encuentran que los profesores de primaria y secundaria utilizan una intensidad alta (entre 72 y 90 dB) un 20% del tiempo, y los de universidad emplean la misma intensidad un 30% del tiempo. En el caso de los profesores de universidad, hay que tener en cuenta que las aulas y espacios en los que imparten clases son más amplios que en otros niveles educativos. Además, el tiempo de fonación durante una clase teórica puede aproximarse al 90% o incluso más en algunas ocasiones (González-Pumariega & Castejón, 2012). Esto supone un tiempo de fonación por hora de habla muy superior al del profesor de educación primaria, pero, en general, tiene un número menor de horas de docencia directa.

Varios grupos de docentes a tener en cuenta son los profesores de educación física, ya que sus clases suelen desarrollarse en espacios abiertos, y los de matemáticas, que suelen usar más la pizarra, aumentando de esta forma la exposición al polvo de tiza o al olor de rotuladores en el caso de no disponer de pizarra digital (Preciado et al., 2005).

Arrizabalaga y Vega (2017) observaron que los profesores que padecían nódulos vocales tenían una media de 25,14 horas de clases semanales y de 4,95 horas de clases diarias (en niveles de educación infantil y primaria principalmente, donde la presencia de nódulos se evidenció mayor). Preciado et al. (2005) no observan diferencias estadísticamente significativas entre el número medio de horas semanales de docencia (media de 20,9 ± 6,42) de los docentes que padecen nódulos y los que no. Este análisis de la carga vocal, no se ha encontrado en niveles educativos universitarios en la literatura consultada. Este hecho puede explicar que la media de horas semanales totales descrita por los autores mencionados sea mayor que la encontrada en el estudio de esta tesis. Concretamente, se ha observado que la media de horas de clase semanales del grupo de profesores es de 6,45 (±2,37), con unos valores de mínimo y máximo de 3 y 12 respectivamente. También se aprecia que el mayor número de horas de dedicación media semanal aparece en el grupo de prevención primaria, con 6,63 horas (±2,50), seguido del grupo de prevención secundaria con una media de 6,50 horas semanales (±2,01), y por último el grupo de prevención terciaria con de 5,40 horas (±2,51). No obstante, no se encuentran diferencias estadísticamente significativas entre las medias. El conocimiento de los perfiles de carga vocal es muy importante para la realización de programas de educación vocal, ya que la distribución de los tiempos de descanso puede resultar determinante en la prevención y rehabilitación de patologías vocales.

Existen distintas opiniones en la literatura consultada acerca de la relación entre el número de años dedicado a la docencia, y la aparición de patologías vocales. Vargas y Muñoz (2017) no encuentran relación entre los años dedicados a la docencia y aspectos relacionados con el desarrollo de una patología vocal. Simberg et al. (2006) recogen que el proceso patológico de una disfonía por el ejercicio de una profesión docente es más frecuente en los 5 primeros años de carrera. En el estudio presentado, la media de años de experiencia docente que se observa en el grupo de profesores es de 20,77 (±10,93), contando con profesores noveles y muy veteranos en un rango de valores desde 0 a 43 años de experiencia. El grupo de profesores que presenta una media mayor de años de dedicación es el asignado a prevención primaria con 21,51 años (±10,91) y un rango entre 2 y 41. El siguiente es el grupo de prevención terciaria, con una media de 21,40 años (±9,92), y un rango entre 12 y 37 años. Por último, el grupo que presenta una media menor que los anteriores es el asignado a prevención secundaria, con una media de 19,24 años (±11,22) y un rango entre 0 y 43. Ranchal y Vaquero (2008) obtuvieron en su estudio una media de 17,25 (± 0,83) años, y un rango de 0 a 41 años en aquellos profesores que presentaban nódulos vocales. Preciado et al. (2005) obtuvieron una media de 19,20 (±9,25) años de dedicación a la docencia, y un rango entre 1 y 41 años. Barbero-Díaz et al. (2010) observaron en su estudio sobre la incapacidad vocal en docentes de la provincia de Huelva una media de años trabajados de 16,76 (±11,07) (rango entre 1 y 42 años).

A pesar de no diferir mucho los datos entre los diferentes autores presentados, no se puede establecer una relación causal entre los años dedicados a la docencia y el desarrollo de problemas vocales, ya que, como se ha indicado anteriormente, existen factores como la carga vocal que pueden influir de forma negativa en la función vocal. De esta forma, sujetos con menor tiempo en la función docente pueden presentar cargas vocales elevadas, hábitos poco saludables y condiciones personales adversas, que pueden provocar actuaciones de mal uso y/o abuso vocal, que otros, con mayor tiempo de dedicación. Este fenómeno, permite sugerir que el tiempo de experiencia en la docencia no define directamente aspectos relacionados con el desarrollo de patologías vocales, aunque unida a otros aspectos negativos pueda favorecerlas.

En este estudio se ha incluido información sobre la percepción subjetiva del padecimiento de un problema vocal, así como el efecto que tiene sobre las esferas emocional, funcional o física. Esto se ha realizado aplicando el cuestionario VHI-10 (Rosen et al., 2004). Existen autores que afirman que las valoraciones que una persona puede realizar acerca de su salud vocal están altamente correlacionadas con sus necesidades vitales (laborales, personales, sociales) (Vilaseca et al., 2004; Vilaseca et al., 2008). La muestra con la que hemos trabajado presenta valores diferentes en cuanto a la puntuación del cuestionario VHI dependiendo del grupo en el que se encuentren. De esta forma, el grupo de profesores tiene una media de 3,53 (±5,02) puntos, con un rango de 0-28. En el grupo de enfermos, la media es 16,36 (±10,82) puntos, con un rango que va desde los 2 hasta los 36 puntos. Y, por último, la puntuación media de los sanos es la más baja de los tres grupos, con 1,68 (±2,70) puntos, y con unos valores mínimo y máximo de 0 y 11, respectivamente. Gassul et al. (2013) consideraron valores totales de VHI-10 menores a 9 como indicadores de voz normal, entre 10-20 como disfonía leve, entre 21-30 disfonía moderada y, finalmente de 31-40 como disfonía grave. Vargas y Muñoz (2017) hallaron un valor medio de puntuación de VHI-10 de 17,63 (±6,08) en una muestra de profesores que realizaron el cuestionario a finales del tercer trimestre del curso académico. Si tenemos en cuenta la correlación que establecen Gassul et al. (2013) entre las distintas puntuaciones que pueden obtenerse en el cuestionario VHI-10 y su correspondencia con niveles de disfonía, se podría decir que en nuestro trabajo las puntuaciones obtenidas en el grupo de enfermos se corresponderían con una disfonía leve en la mayor parte de los sujetos. Sin embargo, debido a que los valores oscilan entre los 2 y 36 puntos, existirían casos que se corresponderían con la clasificación desde disfonía moderada hasta grave, estableciéndose una relación entre la salud vocal percibida y la finalmente diagnosticada (nódulos vocales). Si se lleva este razonamiento al grupo de profesores, se puede decir que la mayor parte de ellos se clasifican en el nivel de voz normal. Sin embargo, al igual que en el caso de los enfermos, la oscilación de los valores de rango (0-28) nos indicaría que podría existir algún sujeto con disfonía moderada, pero que ninguno llegaría a disfonía grave. No obstante, esta escala es subjetiva al igual que los límites propuestos por Gassul et al. (2013).

En nuestro estudio, consideramos importante la valoración subjetiva sobre la incapacidad vocal percibida que realiza el grupo de docentes y analizamos las puntuaciones obtenidas en los distintos grupos de riesgo.

De esta forma, se puede observar que la mayor puntuación del cuestionario VHI-10 aparece en los profesores asignados al nivel de prevención terciaria con una media de 11,20 puntos (±10,23), seguida de los asignados al nivel de prevención secundaria con una media de 5,16 puntos (±5,82) y por último a los profesores del grupo de nivel primario con un valor medio de 2,25 (±3,28). Estos resultados indican que el cuestionario VHI-10 está proporcionando información fiable.

A pesar de que las relaciones de causalidad entre la aparición de un trastorno vocal y el desempeño de una profesión no aparecen reflejadas en muchos estudios epidemiológicos en la actualidad, existen investigaciones llevadas a cabo entre determinados grupos de trabajadores que sugieren que, diferentes trastornos vocales están relacionados con la labor profesional desempeñada (Preciado et al., 2008; Russell et al., 1998). La prevalencia de las patologías vocales continúa siendo un aspecto difícil de concretar en la población general, y también entre los docentes. Muchos de los profesionales dedicados a la enseñanza integran las alteraciones vocales como parte natural de su práctica laboral, y no notifican enfermedad alguna. Existen estudios que estiman que la prevalencia de lesiones vocales en la población general se sitúa entre el 6-15%, mientras que en los docentes llegaría hasta el 90% (Hazlett et al., 2011; Martins et al., 2014; Williams, 2003). Roy et al., en 2004 elaboraron un estudio a partir de una muestra de 2.531 individuos de los estados de Iowa y Utah (Estados Unidos de América). De ellos, el 49,1% eran docentes, y el resto desempeñaba otras profesiones. Tras la realización de un cuestionario telefónico sobre trastornos vocales, determinaron que el 58% de los docentes habían padecido algún problema vocal en algún momento de su carrera, frente a únicamente el 29% de los que ejercían otra profesión. Behlau et al. (2012) realizaron un estudio con una muestra de 3.265 participantes de 27 estados de Brasil, de los que 1.651 eran docentes y 1.614 desempeñaban otras profesiones. Encontraron que el 63% de los docentes estudiados habían padecido problemas vocales en algún momento de su vida. Sodersten et al. (2002) plantean que entre el 32% y el 72% de los profesores presentan problemas vocales. Bovo et al. (2007) recogen que hasta el 50% de los profesores presentan trastornos vocales. En España, Preciado et al. (2005), en un estudio realizado con una muestra de 527 profesores, recogen una prevalencia de trastornos vocales del 57%. En nuestro estudio, la prevalencia observada es del 16,9% de un total de 120 profesores analizados. Se debe tener en cuenta que, en la mayor parte de los estudios, la generalización de los estudios al colectivo docente supone la desestimación de las diferencias entre los diferentes niveles de enseñanza. En nuestro caso, la investigación llevada a cabo se ha centrado en profesores de universidad, donde la prevalencia suele ser menor.

Los nódulos vocales, además de ser la única enfermedad profesional relacionada con el uso profesional de la voz recogida por la legislación española, es la patología orgánica que mayor incidencia y prevalencia tiene entre los docentes. Urritikoetxea et al. (1995) utilizaron una muestra de 1046 docentes para estudiar la prevalencia de lesiones vocales orgánicas. En su investigación, encontraron la presencia de lesiones en 218 de los encuestados. La patología más prevalente fue la de nódulos vocales, con 94 individuos (43,12% de los afectados), seguida del edema de Reinke con 39, hipertrofia de bandas vestibulares con 24, y por último pólipos y otras lesiones menos frecuentes con 19.

En el mismo sentido, Tavares y Martins (2007) observaron que el 63,16% de las lesiones orgánicas encontradas en un estudio realizado a 80 docentes fueron nódulos vocales. Pérez y Preciado (2003) identificaron algunos factores potencialmente predisponentes para el desarrollo de nódulos vocales, como son la patología vocal previa, la microcirugía laríngea, la cirugía nasal y el reflujo gastro-esofágico. En su estudio, inciden en la importancia de la prevención y el diagnóstico de la salud vocal en docentes desde niveles previos al ejercicio de la profesión. De esta forma, lesiones fonotraumáticas como son los nódulos vocales, podrían prevenirse e incluso diagnosticarse en estudiantes de los distintos Grados de Educación. Según el Instituto Nacional de Estadística, en España, el número de casos de nódulos vocales en mujeres fue 12 veces mayor que en el caso de los hombres en el año 2017. Estos datos van en la misma línea que los obtenidos en nuestro trabajo, donde el 84% de los individuos con nódulos vocales encontrados en el grupo de enfermos son mujeres. La determinación del riesgo vocal se ha convertido en una necesidad en el ámbito de la prevención laboral. Disponer de eficaces y sólidos métodos de diagnóstico que puedan ayudar a la detección en estadios iniciales de la enfermedad puede suponer un ahorro de costes económicos (bajas laborales, ingresos hospitalarios por intervenciones quirúrgicas) a la administración, una mejora en la calidad de atención del paciente y un mejor aprovechamiento de los recursos.

En la actualidad existen métodos objetivos de evaluación vocal a partir de análisis computerizados que permiten al especialista verificar lo observado por el paciente y por su propio oído clínico (Fernández et al., 2006b; Señaris et al., 2006). Los avances en la utilización de técnicas acústicas para el diagnóstico de patologías vocales han llevado a la ejecución de un gran número de experimentos con muy diferentes fines. Gassul et al. (2013) utilizan el análisis de parámetros acústicos para estudiar el impacto de un programa de educación de la voz presentado en la asignatura (obligatoria) Educación de la Voz en estudiantes de magisterio de la Facultad de Ciencias de la Educación de la Universidad Autónoma de Barcelona. Gurlekian y Molina (2012) desarrollan y proponen un sistema de evaluación objetiva del riesgo vocal a partir de la medición y análisis de parámetros acústicos para los tres grupos de profesionales de la voz que ellos consideran más afectados: docentes (de cualquier nivel educativo), actores, locutores, cantantes y otros trabajadores de las artes o espectáculo, y telefonistas. Casado et al. (2001) presentan un estudio a través del análisis acústico para cuantificar las diferencias entre sujetos con nódulos y pólipos vocales y sujetos normales. Rufo et al. (2019) proponen un análisis bayesiano de tres características acústicas (Shimmer, HNR, RPDE) y la variable sexo, para determinar el nivel de riesgo de padecer trastornos vocales en profesores universitarios. Este enfoque permite incorporar un resumen integrado de la evidencia disponible y la incertidumbre asociada con la evaluación de las utilidades. Paniagua et al. (2017) proponen la valoración del efecto de la fonomicrocirugía en pacientes con patología vocal como son el edema de Reinke y pólipos vocales, utilizando un análisis acústico con medidas lineales (Jitter) y no lineales (DFA, RPDE, PPE) antes y después de la cirugía. En esta misma línea, Pérez et al. (2019) proponen el uso de características acústicas lineales (Jitter, Shimmer, HNR, GNE) y no lineales (ZCR (Zero Crossing Rate), D2, LZ2 (Lempel Complexity), FMMI (First Mínimum in Mutual Information) y FZCF (First Zero in Correlation Function)) para evaluar la recuperación completa de los sujetos con edema de Reinke después de la fonocirugía.

En el trabajo que se presenta se muestra el diseño de un programa de prevención de nódulos vocales (o patologías de la voz similares) llevado a cabo tras un primer diagnóstico (automático) realizado con un método de clasificación basado en el reconocimiento de patrones. Se han utilizado tres características lineales basadas en la perturbación de frecuencia, perturbación de amplitud y el ruido glótico (Jitter relativo, Shimmer local y GNE), y siete no lineales (PermutEn, FuzzyEn, SampleEn, RPDE, DFA y PPE). Las características no lineales son de particular relevancia para la práctica clínica, ya que han demostrado tener una buena capacidad predictiva en el diagnóstico de patologías vocales asociadas a disfonías importantes (Jiang et al., 2009; Paniagua et al., 2018). Estas enfermedades, debido a los patrones de vibración vocal que producen, pueden presentar fenómenos no lineales y aleatorios, con modulaciones y bifurcaciones (señales Tipo 2), y ciclos irregulares y aperiódicos (señales Tipo 3) (Titze, 1995) en mayor medida que las voces sanas, que se aproximan más al modelo lineal de filtro fuente.

A pesar de la alta sensibilidad que tiene Jitter para afecciones en las que se producen cambios o asimetrías en la masa de las cuerdas vocales (aumentando mucho sus valores) (Koike, 1973), hay que ser cautos al relacionar valores patológicos de Jitter con una determinada afección. Debido a la relación existente entre la variable sexo y las características de una voz humana, en la investigación presentada se ha optado por la variable de Jitter relativo, que mide la variación interciclos en porcentaje, y no se ve afectada por la frecuencia fundamental, ni sufre variaciones según el sexo. Los valores obtenidos en los diferentes grupos analizados en nuestro experimento muestran que, para el grupo de enfermos, el valor medio de Jitter relativo obtenido es de 0,98 (±1,26), mayor que el encontrado en el grupo de profesores (0,54 ± 0,32) y en el grupo de sanos (0,43 ±0,18). Jiang et al. (2009) obtienen valores similares a los nuestros para el grupo de voces sanas (0,40±0,19) en su estudio de caracterización de voces con nódulos, pólipos y sanas. Sin embargo, los valores para el grupo de pacientes con nódulos vocales son menores a los nuestros, con un valor medio de 0,59 (±0,23). Scalassara et al. (2009) obtienen valores superiores a los nuestros en el grupo de sanos (0,52 ±0,40) en su estudio para describir las características acústicas de voces sana, con nódulos vocales y edema de Reinke. También obtienen resultados superiores en el caso de voces con nódulos vocales, donde el valor medio de Jitter se sitúa en 3,05 (±3,13). Gurlekian y Molina (2012) dividen una muestra de 84 docentes en tres grupos de voces: normales, con riesgo vocal y con alteración vocal. En el primer grupo, obtienen un valor medio de Jitter de 0,81 (±0,29), superior al nuestro en el grupo de sanos. No se ha podido comparar el valor obtenido en el grupo de alteraciones de la voz, dado que no se especifica la patología. Casado et al. (2001) encuentran valores medios de Jitter en sujetos sanos de 0,24 (±0,14) y de 0,35 (±0,08) en sujetos con nódulos. Cobeta y Núñez (2013) definen el valor medio de Jitter relativo en condiciones de normalidad en 0,97. Gurlekian y Molina (2012) establecen que la escala de Jitter relativo oscila entre el 0,1 y 1 en el segmento de la normalidad, y entre el 1 y el 10 en el de la alteración vocal. Atendiendo a estas consideraciones y conociendo que valores de Jitter de 1% son percibidos como disfonía con voz rugosa, podemos sugerir que nuestros resultados en los distintos grupos pueden ser considerados de forma objetiva.

El parámetro Shimmer mide la variabilidad de la amplitud ciclo a ciclo, y representa una medida de la estabilidad de la fonación. Se relaciona de modo inverso con la intensidad vocal, disminuyendo al aumentar la intensidad. Pequeñas variaciones ciclo a ciclo pueden considerarse normales durante la fonación, pero una variabilidad excesiva indica una posible patología. Para nuestro estudio se ha utilizado Shimmer local, que no presenta diferencias por sexo.

El valor medio de Shimmer local encontrado en los distintos grupos analizados de nuestra investigación es similar para el grupo de sanos y profesores (0,04 con desviaciones ± 0,01 y ± 0,02 respectivamente), y mayor en el de enfermos (0,06 ±0,03). Estos valores son inferiores a los encontrados por Casado et al. (2001) en el grupo de voces sanas (2,10 ±0,14), y en el de voces afectadas por nódulos vocales (3,25 ±0,86). Gurlekian y Molina (2012) encuentran también resultados superiores a los nuestros en el caso de voces sanas, con un valor medio de Shimmer de 0,21 (±0,07). Jiang et al. (2009) obtienen un valor medio de Shimmer para voces normales de 0,91 (±0,72), y de 1,28 (±0,69) para voces con nódulos vocales. Al igual que en el caso del parámetro anterior (Jitter), los resultados obtenidos en los estudios presentados muestran diferencias atribuibles también a las condiciones en las que se han realizado las grabaciones, así como a los equipos utilizados. Además, hay que tener en cuenta, que el parámetro Shimmer es una característica que presenta gran variabilidad entre los distintos individuos (Cobeta & Núñez, 2013).

Como puede apreciarse existe una disparidad de resultados para caracterizar grupos de voces aparentemente con las mismas condiciones de salud o patología. Esto puede deberse entre otras cosas, a que los métodos de extracción y de análisis utilizados por los distintos autores no son los mismos, y tampoco lo son las condiciones ambientales para la realización de grabaciones acústicas.

Las características Jitter y Shimmer son las más utilizadas en la literatura científica y ambas son lineales. En esta tesis se parte de la hipótesis de que la incorporación de medidas no lineales como biomarcadores puede incrementar la eficiencia de la discriminación entre voces sanas y patológicas.

Hasta donde llega nuestro conocimiento, existe un déficit de programas de prevención para profesores basados en biomarcadores acústicos. Se han desarrollado protocolos de detección de disfonías por el uso profesional de la voz, como el llevado a cabo en Navarra por Arrizabalaga y Vega (2017), o programas basados en niveles de actuación en la prevención de disfonías en docentes, como el establecido por Castejón (2014), en el que contemplaba tres niveles de actuación en la prevención de las disfonías (nivel conceptual y actitudinal, nivel procedimental y nivel ecológico). Sin embargo, no se han implementado programas en los que se utilice el análisis de características acústicas para establecer niveles de riesgo de diagnóstico asociados a actuaciones de prevención.

Para el modelo de clasificación utilizado en el experimento, el subconjunto de características seleccionado consistió en GNE y tres medidas no lineales (no se seleccionaron medidas de perturbación convencionales).

GNE compara la cantidad de excitación de la voz causada por las oscilaciones de las cuerdas vocales, con la excitación generada por el ruido (turbulencias) (Michaelis et al., 1997). PPE es una medida no lineal de variación de frecuencia fundamental.

Su gran ventaja en comparación con las medidas de perturbación clásicas (variantes de fluctuación de fase) es que puede distinguir variaciones naturales de frecuencia fundamental provenientes de un vibrato suave natural de variaciones disfónicas producidas por un trastorno de la voz. Este parámetro fue introducido por Little et al., en 2009. Finalmente, el proceso de selección de características también incluyó D2 y SampleEn, dos características no lineales complementarias que caracterizan la complejidad de un sistema dinámico bajo dos perspectivas diferentes. Por un lado, midiendo la dimensionalidad de la dinámica subyacente (D2), y por otro, estimando la imprevisibilidad dentro de las series de tiempo (SampleEn). Como es esperable, el método de asignación automática cometió algunos errores. Específicamente, 6 de los 103 profesores que informaron que no tenían problemas de voz fueron clasificados en el grupo de alto riesgo sin pertenecer a este grupo tras su análisis, y 3 de los 17 profesores que reconocieron la existencia de problemas de voz, no fueron asignados al grupo de riesgo correcto. Esto implica que únicamente el 7,5% de las asignaciones debió ser corregido por el profesional de medicina ocupacional. Se encontró una asociación significativa entre la existencia de problemas de voz y el nivel de prevención asignado. Por lo tanto, en base a los resultados obtenidos en este experimento, consideramos que el enfoque metodológico que se aplicó es prometedor.

Además, el método propuesto para la asignación de profesores a grupos de riesgo de padecer nódulos vocales (u otras patologías con similares características fonatorias) es no invasivo y de bajo coste. La grabación de las voces se realizó con equipos asumibles económicamente (ordenador, tarjeta de sonido y micrófono) y la asignación de grupo se realizó con software propio desarrollado específicamente para esta tarea. La parte más costosa del programa preventivo es la relacionada con las acciones tomadas en el nivel de prevención terciaria. Esto incluye visitas al otorrinolaringólogo y seguimiento por un logopeda. El programa de prevención primaria tiene un coste muy reducido, ya que el curso ofrecido se puede impartir en grupos grandes y, además, se podría publicar en línea, reduciendo aún más los costes. El programa de prevención secundaria está dirigido a pequeños grupos de hasta diez personas. No se ha realizado un análisis de costes específico pero estas acciones se integrarían dentro del Programa de Prevención de Enfermedades y Promoción de la Salud del Servicio de Prevención de Riesgos. Este programa cubre muchos otros riesgos laborales y el coste económico exacto del programa de prevención vocal en sí es difícil de cuantificar. Con el presente experimento, no es posible saber exactamente cuántos sujetos solicitarían una baja laboral si no asistieran al nivel de prevención correspondiente. Sin embargo, se sabe que, en promedio, cinco días de baja por enfermedad pueden cubrir el coste económico del tratamiento completo necesario para un sujeto en prevención terciaria durante un mes.

Esta tendencia de actuaciones basada en la eficiencia de procedimientos es el común denominador de la medicina basada en la evidencia, que tiene como uno de sus objetivos la ayuda en la toma de decisiones clínicas, la mejora de la calidad de la atención del paciente, la educación del paciente, la reducción de los costes, el aprovechamiento de los recursos y la reducción de riesgos también desde el punto de vista legal (Biddle et al., 2002). Por lo tanto, y sumando a todo esto el obvio bienestar personal que puede producir el tratamiento y mejora de este tipo de patologías, se espera que el programa de prevención sea muy favorable en todos los aspectos contemplados.

Por último, es importante destacar que, a pesar de las prometedoras posibilidades de este experimento, existe la limitación de sólo poder llegar a aquellos sujetos que se realizan el reconocimiento médico anual en el Servicio de Prevención de la Universidad de Extremadura. Teniendo en cuenta la voluntariedad de esta revisión, no es posible abarcar, y por lo tanto aplicar, el programa de prevención a todo el colectivo docente actualmente en activo.

Capítulo 7. Conclusión

Los profesionales de la voz son propensos a desarrollar trastornos vocales debido al uso continuo y sostenido de la voz en su entorno laboral. Aunque estos trastornos afectan a diferentes perfiles, ser mujer, profesora y estar en la cuarta década de la vida supone un riesgo mayor de padecer patologías vocales como, por ejemplo, los nódulos vocales. El estilo de vida, los aspectos emocionales y psicológicos, los hábitos de higiene vocal y la técnica vocal de cada persona son factores relacionados con la salud vocal que deben tenerse en cuenta. No obstante, en el desarrollo de una patología vocal, el ambiente de trabajo, el tiempo de trabajo, el tiempo de descanso o los espacios en los que desarrolla la actividad profesional son aspectos que influyen en el desarrollo de un trastorno vocal.

En la literatura científica existe un déficit de programas preventivos de atención a la salud vocal de los profesionales de la voz en general, y de los profesores en particular. Esto hace que sea necesario planificar, desarrollar e implementar programas preventivos que permitan identificar riesgos y proporcionar la aplicación rápida de medidas preventivas y/o correctivas. En esta tesis doctoral se demuestra que es posible definir y aplicar un programa preventivo de trastornos vocales basado en biomarcadores acústicos a profesores de universidad.

La idea de utilizar modelos de clasificación con las características acústicas extraídas de grabaciones de voz para asignar a los profesores a tres grupos de riesgo vocal es innovadora. Diversos modelos de clasificación con una selección de los mejores predictores han proporcionado buenos resultados. En concreto, los mejores porcentajes de clasificación correcta se han obtenido con un modelo basado en árboles de decisión denominado Random Forest, que ha proporcionado valores de porcentaje de acierto, sensibilidad y especificad del 92%. Esto se ha conseguido utilizando tanto características acústicas lineales como no lineales.

A partir de los grupos de riesgo asignados automáticamente, los profesores se integran en tres niveles de prevención, que conllevan acciones encaminadas a la protección, promoción, prevención y rehabilitación de la salud vocal. Los diferentes niveles de prevención se han propuesto específicamente para la difusión de una educación sanitaria que estimule la promoción y el mantenimiento de la salud vocal, así como la identificación temprana de sujetos asintomáticos y la aplicación de medidas correctivas cuando la enfermedad está presente. Hasta donde llega el conocimiento de la doctoranda, este es el primer programa preventivo basado en biomarcadores acústicos que contiene todos los elementos necesarios desde el inicio del proceso de captación de profesores hasta la rehabilitación.

Aunque el experimento se ha desarrollado con la patología de nódulos en las cuerdas vocales y trastornos acústicamente compatibles con esta, la metodología es extensible a otras patologías que, por su prevalencia, características de desarrollo, evolución y tratamiento también requieren una atención especial como pólipos vocales, edemas de Reinke o quistes.

La selección de los nódulos vocales para probar esta metodología se ha realizado porque constituyen el trastorno de la voz más común entre el personal docente y representan la única enfermedad profesional relacionada con el uso profesional de la voz recogida en la legislación española. Del mismo modo, la metodología propuesta es extensible a otros colectivos que utilicen la voz como herramienta de trabajo (teleoperadores, cantantes, actores, locutores, entrenadores).

Finalmente, la implementación de este modelo de prevención en una aplicación para teléfonos inteligentes podría proporcionar a los profesionales de la voz y a los usuarios ventajas considerables en términos de ahorro de tiempo y dinero y lo que es más importante, de mantenimiento o mejora de la salud. Sin embargo, aún falta trabajo por realizar para la implementación de un modo generalizado, ya que los diferentes dispositivos de grabación y el ruido ambiental tienen efectos importantes en las características acústicas que deben extraerse. Un análisis extensivo de robustez de las características acústicas es necesario y se plantea como trabajo futuro.

Referencias

Al-Nasheri, A., Muhammad, G., Alsulaiman, M., Ali, Z., Mesallam, T. A., Farahat, M., Malki, K.H., & Bencherif, M. A. (2017). An investigation of multidimensional voice program parameters in three different databases for voice pathology detection and classification. *Journal of Voice*, 31(1), 113-e9-113-e18.

Altman, K. W. (2007). Vocal fold masses. *Otolaryngologic Clinics of North America*, 40(5), 1091-1108.

Aronson, A.E. (1990). *Clinical voice disorders. An interdisciplinary approach*. New York: Thieme Inc.

Arrizabalaga, S., & García, V. (2017). Evaluación de la notificación de disfonías en personas con uso profesional de la voz como suceso centinela en Navarra. Años 2013-2015. *Revista Española de Salud Pública*, 91, 1-11

Avilés, F.J., Doménech, E., Figuerola i Massana, E. (2014). Patología de la voz hablada y de la voz cantada. *Libro virtual de formación en ORL*, 117, 1-20. Consultado el 12 de febrero de 2018 desde http://seorl.net/PDF/Laringe%20arbor%20traqueo-bronquial/117%20-%20PATOLOG%C3%8DA%20DE%20LA%20VOZ%20HABLADA%20Y%20DE%20LA%20VOZ%20CANTADA.pdf

Baghai-Ravary, L., & Beet, S. W. (2012). Automatic speech signal analysis for clinical diagnosis and assessment of speech disorders. Oxford: Springer Science & Business Media.

Baken, R. J., & Orlikoff, R. F. (2000). *Clinical measurement of speech and voice*. San Diego: Singular Thomson Learning.

Barbero-Díaz, F. J., Ruiz-Frutos, C., Del Barrio, A., Bejarano, E., & Alarcón, A. (2010). Incapacidad vocal en docentes de la provincia de Huelva. *Medicina y seguridad del trabajo*, 56(218), 39-48.

Behlau, M., Zambon, F., Guerrieri, A. C., & Roy, N. (2012). Epidemiology of voice disorders in teachers and nonteachers in Brazil: prevalence and adverse effects. *Journal of voice*, 26(5), 665-e9.

Belafsky, P.C., Postma, G., Reulbach, T.R., Holland, B. W., Koufman, J. A. (2002). Muscle tension dysphonia as a sign of underlying glottal insufficiency. *Otolaryngol Head Neck Surg*., 127, 448-51.

Biddle A, Watson L, & Hooper C. (2002). Using evidence-based research in speech-language pathology. A paper presented to the meeting of the American Speech, Language, and Hearing Association, Atlanta.

Blomgren, M., Chen, Y., Ng, M. L., & Gilbert, H. R. (1998). Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers. *J Acoust Soc Am*, 103(5), 2649-2658.

Bonilla, L. N. T., Pérez, M. C., & Camacho, J. (2018). Patología benigna de laringe en docentes y cantantes. *Acta de Otorrinolaringología & Cirugía de Cabeza y Cuello,* 39(1), 15-23.

Boone, D.R. (1983) *Voice and voice therapy* (2nd ed.). London: Prentice-Hall.

Bovo, R., Galceran, M., Petruccelli, J., & Hatzopoulos, S. (2007). Vocal problems among teachers: evaluation of a preventive voice program. *Journal of voice*, 21(6), 705-722.

Bravi, A., Longtin, A., & Seely, A. J. (2011). Review and classification of variability analysis techniques with clinical applications. *Biomedical engineering online*, 10(1), 90.

Buekers, R., Bierens, E., Kingma, H. y Marres, E.H. (1995). Vocal load as measured by the voice accumulator. *Folia Phoniatrica y Logopaedica*, 47, 247-302.

Bustos, I. (1991). Reeducación de problemas de la voz. 4ª ed. Madrid: Cepe.

Calle-Alonso, F., Pérez, C. J., Arias-Nicolás, J. P., & Martín, J. (2013). Computer-aided diagnosis system: A Bayesian hybrid classification method. *Computer methods and programs in biomedicine,* 112(1), 104-113.

Calvache Mora, C. A. (2015). Objetividad de la electroglotografía. Aplicaciones clínicas e investigativas en la voz. *Revista de Logopedia, Foniatría y Audiología*, 35(3), 134-142.

Cantor, L.C, Vogel I., Burdorf, A. (2013). Voice disorders in teachers and their associations with work-related factors: a systematic review. *J Commun Disord*, 46(2), 143-55.

Casado, J. C., Adrián, J.A., Conde, M., Piédrola, D., Povedano, V., Muñoz, E., …Jurado, A. (2001). Estudio objetivo de la voz en población normal y en la disfonía por nódulos y pólipos vocales. *Acta Otorrinolaringologica Española*, 52, 476-482.

Castejón, L (2013). Prevención de disfonías funcionales en docentes: el papel de los profesores de universidad con estudiantes del grado de Maestro. *Magister*, 25 (1), 67-73.

Castejón, L. (2014). Prevención de disfonías funcionales en el profesorado universitario: tres niveles de acción preventiva. *Aula abierta*, 42(1), 9-14.

Cervera, F., Vega, F., García Tapia, R. (1996). Lesiones benignas de las cuerdas vocales. En: García-Tapia, R. & Cobeta, I. (Eds) *Diagnóstico y Tratamiento de los Trastornos de la Voz* (pp.223-241), Madrid: Editorial Garsi.

Czerwonka L, Jiang JJ, Tao, C. (2008) Vocal nodules and edema may be due to vibration-induced rises in capillary presure. *Laryngoscope* (118), 748-52.

Chuang Ch, Á., Badía, P., & Napolitano, C. (2017). Quistes de cuerda vocal: Experiencia en 44 pacientes del Centro de Voz del Departamento de Otorrinolaringología de la Pontificia Universidad Católica de Chile. *Revista de otorrinolaringología y cirugía de cabeza y cuello,* 77(3), 239-245.

Cobeta, I., García, I., Lara, I. (2008). Disfonías funcionales. En Suárez, C. et al., *Tratado de otorrinolaringología y cirugía de cabeza y cuello,* 2ªed (pp.2643-2654). Madrid: Editorial Médica Panamericana.

Cobeta, I., Lara, A., García, I. (2008b). Evaluación clínica de la fonación. Laboratorio de voz. En: Suarez C, et al. (Eds). *Tratado de otorrinolaringología y cirugía de cabeza y cuello*. 2ª ed. (pp 2545-2567), Madrid: Editorial Medica Panamericana.

Cobeta, I. & Núñez, F. (2013). Laboratorio de voz. Análisis de la señal acústica. En Cobeta, I., Núñez, F. & Fernández, S (Eds). *Patología de la voz* (pp.188-198), Barcelona: Marge Médica Books

Cobeta, I., Núñez, F. & Fernández, S. (2013b). Exploración funcional por la imagen. Estroboscopia. En Cobeta, F., Núñez, I. & Fernández, S (Eds). *Patología de la voz* (pp.146-158), Barcelona: Marge Médica Books.

Cobeta, I., Núñez, F. & Fernández, S. (2013c). Voz normal y clasificación de disfonías. En Cobeta, F., Núñez, I. & Fernández, S (Eds). *Patología de la voz* (pp.237-241), Barcelona: Marge Médica Books.

Coll, R. (2013). Higiene vocal. En Cobeta, F., Núñez, I. & Fernández, S (Eds*). Patología de la voz* (pp.483-487), Barcelona: Marge Médica Books

Colton, R. H., Casper, J. K., & Leonard, R. (2006). Differential diagnosis of voice problems. En *Understanding voice problems. A physiological perspective for diagnosis and treatment*. (3rd ed) (pp. 12-63). United States of America: Lippincott Williams & Wilkins, 12-63.

Colton, R. H., Casper, J. K., & Leonard, R. J. (2011). Understanding voice problem: A physiological perspective for diagnosis and treatment: Fourth edition. Baltimore, United States of America: Lippincott Williams & Wilkins.

Cuestas, G., Urquizo, M., Demarchi, V., Zanetta, A., Lobos, P., & Razetti, J. (2013). Tratamiento de las estenosis laríngeas adquiridas en pediatría: Serie de casos. *Archivos argentinos de pediatría*, 111(6), e136-e140.

Dejonckere, P. H., & Kob, M. (2009). Pathogenesis of vocal fold nodules: new insights from a modelling approach. *Folia Phoniatrica et Logopaedica*, 61(3), 171-179.

De Oliveira Bastos, P. R. H., & Hermes, E. C. (2018). Effectiveness of the Teacher's Vocal Health Program (TVHP) in the Municipal Education Network of Campo Grande, MS. *Journal of Voice*, 32(6), 681-688.

Díaz de Cerio, P., Gascón M. C. & Lacosta, J. L. (2010). Pólipos, Nódulos, Quistes, Edema de Reinke y Laringoceles. Microcirugía laríngea. *Libro virtual de formación en ORL*. Complejo Hospitalario San Millán- San Pedro. Logroño. Consultado el 10 de octubre de 2018 desde http://seorl.net/PDF/Laringe%20arbor%20traqueo-bronquial/108%20-%20P%C3%93LIPOS,%20N%C3%93DULOS,%20QUISTES,%20EDEMA%20DE%20REINKE%20Y%20LARINGOCELES.%20MICROCIRUG%C3%8DA%20LAR%C3%8DNGEA.pdf

Dikkers, F. G., & Nikkels, P. G. (1995). Benign lesions of the vocal folds: histopathology and phonotrauma. *Annals of Otology, Rhinology & Laryngology*, 104(9), 698-703.

Dodero A, Hortas MA, Wilder F. (2005). Metodología de estudio de las alteraciones de la voz y baremos de incapacidad vocal. *Cuadernos de medicina forense*, 4(1), 41-51.

Dosal González, R (2014). *Producción de la voz y el habla. La fonación* (trabajo fin de grado). Escuela de Enfermería Casa Salud Valdecilla, Universidad de Cantabria, Santander, España.

Duffy, J. R. (2005). *Motor Speech Disorders: Substrates. Differential Diagnosis and Management, (2ªed)* (p592). New York: Mosby.

Elhendi, H.W. (2005). *Medidas objetivas y subjetivas para la evaluación de los resultados tras el tratamiento rehabilitador de los nódulos vocales y las disfonías funcionales*. (Tesis doctoral). Universidad de Santiago de Compostela. Santiago de Compostela (España).

Elhendi, W., Caravaca, A., & Santos, S. (2012). Estudio epidemiológico de pacientes con disfonías funcionales. *Anales de Otorrinolaringología Mexicana*, 57(1), 44-50).

Elhendi Halawa, W., Rodríguez Fernández-Freire, A., & Santos Pérez, S. (2012b). Estudio epidemiológico de los pacientes con nódulos vocales*. Revista de Logopedia, Foniatría y Audiología,* 32(4), 164-170.

Elhendi, W., Vázquez, I., & Santos, S. (2012c). Medición de la discapacidad vocal en pacientes con nódulos vocales. *Revista Investigación en Logopedia*, 2, 1-14

Estellés, M.T. (2005) *Prevención en el deterioro vocal de las personas mayores. Manifestaciones clínicas y programa de tratamiento*. En: La voz en nuestros mayores. 6º Simposium Nacional de Logopedas de España A.L.E. (42-51). Valencia.

Farías, P. (2007). *Ejercicios que restauran la función vocal: observaciones clínicas*. Buenos Aires: Librería Akadia.

Farías, P. (2013). Voz hablada. En Cobeta, F., Núñez, I. & Fernández, S (Eds). *Patología de la voz* (pp.500-515), Barcelona: Marge Médica Books

Farias, P. (2014). Mejoría del cierre glótico y de la onda mucosa con terapia en Sulcus Vergeture: Reporte de un caso. *Revista Areté*,14 (1), 6-19.

Farias, P. (2016). *Guía clínica para el especialista en laringe y voz* (1ª Ed). Buenos Aires: Librería Akadia.

Fazio, S., Ortega, A. G., & Sáenz, A. (2014). Disfonías crónicas en adultos. *Revista Médica Universitaria,* 10(1),1-25.

Fernández, R. L., Damborenea, D. T., Rueda, P. G., Leache, J. P., del Alamo Campos, M. A., Llorente, E. A., & Naya, M. G. (1999). Acoustic analysis of the normal voice in nonsmoking adults. *Acta Otorrinolaringológica Española*, 50(2), 134-141.

Fernandez, C. P., & Preciado López, J. (2003). Nódulos de cuerdas vocales. Factores de riesgo en los docentes. Estudio de casos y controles. *Acta Otorrinolaringológica Española*, 54(4), 253-260.

Fernández, S., Vázquez de la Iglesia, F., Marqués, M., & García-Tapia, R. (2006). La historia de la voz. *Rev Med Univ Navarra*, 50(3), 9-13.

Fernández, S., Marques, M., De la Iglesia, F., & Gimeno, C. (2006b). Discapacidad vocal. *Rev Med Univ Navarra*, 50(3), 73-80.

Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems?. *The Journal of Machine Learning Research*, 15(1), 3133-3181.

Ferrein, A. (1741). De la formation de la voix de l'homme. *Hist Acad Roy Sci*, 3, 409-432.

Ferrein, A. (1754). *Mémoire de l'Académie royale des Sciences, séance du 15 novembre 1741.* París.

Ford, C. N., Bless, D. M., Campos, G., & Leddy, M. (1994). Anterior commissure microwebs associated with vocal nodules: detection, prevalence, and significance. *The Laryngoscope*, 104(11), 1369-1375.

Ford, C. N., Inagi, K., Khidr, A., Bless, D. M., & Gilchrist, K. W. (1996). Sulcus vocalis: a rational analytical approach to diagnosis and management. *Annals of otology, rhinology & laryngology*, 105(3), 189-200.

Ford, C. N. (1999). Advances and refinements in phonosurgery. *The Laryngoscope*, 109(12), 1891-1900.

Freeman, M. & Fawcus, M. (2000). *Voice disorders and their management* (3ª ed). Springer. London: Whurr Publishers Ltd.

Frumento, A.S. (1995). Elementos de biofísica de la audición. En Frumento, A.S. *Biofísica (*3ªed) (pp. 377-381). Madrid, España: Mosby-Doyma

Gamero, A. (2013). "Extracción de características de voz para la clasificación de individuos", Trabajo Fin de Grado en Ingeniería de Telecomunicación. Escuela Politécnica, Universidad de Extremadura.

Gañet Benavente R, Serrano Estrada C, Gallego Pulgarín I (2007). Patología vocal en trabajadores docentes: influencia de factores laborales y extralaborales*. Arch Prev Riesgos Labor*, 10 (1), 12-17.

García, R. H., Neves, E. R. B., Bosque, C., & Mendes, E. L. (2014). Voice disorders in teachers. A review. *Journal of voice*, 28(6), 716-724.

García-Río, F., Calle, M., Burgos, F., Casan, P., Del Campo, F., Galdiz, J.B., … Puente-Maestu, L. (2013). *Arch Bronconeumol*, 49(09), 388-401.

García-Tapia, R. (1996*).* Antecedentes Históricos. En García-Tapia, R. & Cobeta, I. (Eds) *Diagnóstico y tratamiento de los trastornos de la voz*. Madrid: GARSI S.A.

García-Tapia R, Fernández S. (1996) Fisiología de la fonación. En: García-Tapia R, Cobeta I. (Eds). *Diagnóstico y tratamiento de los trastornos de la voz*. Ponencia Oficial de la Sociedad Española de ORL y Patología Cervico-Facial (pp.54-86). Madrid: Garsi.

García-Tapia, R. & Fernández, S. (2013). Antecedentes Históricos. En Cobeta, F., Núñez, I. & Fernández, S (Eds). *Patología de la voz* (pp. 23-28), Barcelona: Marge Médica Books

Garrido, M., Castro, R., Málaga, J., & Gorospe, J. M (2004). Prevención de los trastornos de la voz en docentes. En: Vázquez, M. C (Ed.) *Trastornos del lenguaje oral* (pp. 107-131). Universidad del País Vasco. Servicio de Publicaciones.

Gassull, C., Godall, P., & Martín, P. (2013). Incidencia de un programa de educación de la voz para futuros docentes en la mejora de parámetros acústicos y perceptivos de la voz. *Revista de Logopedia, Foniatría y Audiología,* 33(1), 8-12.

Giovanni, A., Lagier, A., & Remacle, M. (2010). Fonomicrocirugía de los tumores benignos de las cuerdas vocales. *EMC-Cirugía Otorrinolaringológica y Cervicofacial*, 11(1), 1-15.

Goller, B (2014). *Espacios Sónicos. Intersecciones entre arquitectura y sonido* (Tesis doctoral). Universidad Politécnica de Madrid, Madrid (España).

González, S. F., San Miguel, D. R., Girbau, M. M., Sarraqueta, L. (2006). Voz del anciano. *Rev Med Univ Navarra,* 50(3), 44-48.

Gotaas, C., & Starr, C. D. (1993). Vocal fatigue among teachers. *Folia Phoniatrica et Logopaedica*, 45(3), 120-129.

Gould, W. J., Sataloff, R. T., & Spiegel, J. R. (1993). *Voice surgery*. St. Louis: Mosby.

Gracida Olvera, G., & Orduña Bustamante, F. (2011). Evocanto: Computer Program for Analyzing the Singing Voice using Digital Signal Processing. *Computación y sistemas*, 15(1), 39-50.

Gurlekian, J. A., & Molina, N. (2012). Índice de perturbación, de precisión vocal y de grado de aprovechamiento de energía para la evaluación del riesgo vocal. *Revista de logopedia, foniatría y audiología,* 32(4), 156-163.

Hazlett, D. E., Duffy, O. M., & Moorhead, S. A. (2011). Review of the impact of voice training on the vocal quality of professional voice users: implications for vocal health and recommendations for further research. *Journal of voice*, 25(2), 181-191.

Henríquez, P., Alonso, J. B., Ferrer, M. A., Travieso, C. M., Godino-Llorente, J. I., & Díaz-de-María, F. (2009). Characterization of healthy and pathological voice through measures based on nonlinear dynamics. *IEEE transactions on audio, speech, and language processing*, 17(6), 1186-1195.

Herrington-Hall, B. L., Lee, L., Stemple, J. C., Niemi, K. R., & McHone, M. M. (1988). Description of laryngeal pathologies by age, sex, and occupation in a treatment-seeking sample. *Journal of Speech and Hearing Disorders*, 53(1), 57-64.

Heuillet-Martin, G., Garson-Bavard, H., & Legré, A. (2007). *Une voix pour tous: Tome 1–La voix normale et comment l'optimiser* (Vol. 1). Marseille, Solal: Groupe de Boeck.

Hirano, M. (1974). Morphological structure of the vocal cord as a vibrator and its variations. *Folia Phoniatrica et Logopaedica*, 26(2), 89-94.

Hirano, M. (1981). Clinical examination of voice. *Disorders of human communication*, 5, 1-99.

Hirano, M. (1989). Objective evaluation of the human voice: clinical aspects. *Folia Phoniatrica et Logopaedica,* 41(2-3), 89-144.

Husson, R. (1950). *Etude des phénomènes physiologiques et acoustiques fondamentaux de la voix chantée.* París: Editions de la revue scientifique.

Husson, R. (1962). *Physiologie de la phonation*. París: Masson.

Jackson-Menaldi, M. C. (1992). *La voz normal*. Buenos Aires: Médica Panamericana.

Jackson-Menaldi MC (2002). *La voz patológica*. Buenos Aires: Editorial Médica Panamericana.

Jackson-Menaldi M.C & Núñez, F. (2013). Valoración de la eficiencia vocal (tiempo de fonación, índice s/e, volúmenes, escalas, fonetograma. En Cobeta, I., Núñez, F. & Fernández, S (Eds). *Patología de la voz* (pp.119-134), Barcelona: Marge Médica Books.

Jacobson, B. H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, M. S., & Newman, C. W. (1997). The voice handicap index (VHI) development and validation. *American Journal of Speech-Language Pathology*, 6(3), 66-70.

Jiang, J. J., Kadota, Y., Yumoto, E., Kurokawa, H., Lin, S. J., & Hanson, D. G. (1998). Quantitative measurement of mucosal wave by high-speed photography in excised larynges. *Annals of Otology, Rhinology & Laryngology*, 107(2), 98-103.

Jiang, J., Lin, E., & Hanson, D. G. (2000). Vocal fold physiology. *Otolaryngologic Clinics of North America*, 33(4), 699-718.

Jiang, J. J., & Zhang, Y. (2002). Chaotic vibration induced by turbulent noise in a two-mass model of vocal folds. *J Acoust. Soc. Amer*., 112(5), 2127-2133.

Johns, M. M. (2003). Update on the etiology, diagnosis, and treatment of vocal fold nodules, polyps, and cysts. *Current opinion in otolaryngology & head and neck surgery*, 11(6), 456-461.

Johnson W, Brown SF, Curtis JF, Edney C.W, Keaster, J. (1965). *Speech handicapped school children*. New York: Harper & Brothers.

Karkos, P. D., & McCormick, M. (2009). The etiology of vocal fold nodules in adults. *Current opinion in otolaryngology & head and neck surgery*, 17(6), 420-423.

Karnell, M. P., Hall, K. D., & Landahl, K. L. (1995). Comparison of fundamental frequency and perturbation measurements among three analysis systems. *Journal of Voice*, 9(4), 383-393.

Kleinsasser, O. (1986). Microlaryngoscopic and Histologic Appearances of Polyps, Nodules, Cysts, Reinkes Edema, and Granulomas of the Vocal Cords. *Vocal Folds Histopathology,* 51-55.

Koike, Y (1973). Application of some acoustic measures for the evaluation of laryngeal dysfunction. *Studia Phonologica*, 7, 17-23.

Koufman, J. A., & Blalock, P. D. (1988). Vocal fatigue and dysphonia in the professional voice user: Bogart-Bacall syndrome. *The Laryngoscope*, 98(5), 493-498.

Koufman J.A., Blalock P.D. (1991) Functional voice disorders. *Otolaryngol Clin North Am*, 24, 1059-73.

Koufman, J. A., & Belafsky, P. C. (2001). Unilateral or localized Reinke's edema (pseudocyst) as a manifestation of vocal fold paresis: the paresis podule. *The Laryngoscope*, 111(4), 576-580.

Kuhn, M. (2015). *Caret: classification and regression training*.

Kyriakou, K., Petinou, K., & Phinikettos, I. (2018). Risk factors for voice disorders in university professors in Cyprus. *Journal of Voice*, 32(5), 643-e1.

Lafon, J. C., & Cornut, G. (1960). Etude de la formation impulsionnelle de la voix et de la parole. *Folia Phoniatrica et Logopaedica*, 12(3), 176-188.

Le Huche, F. (1982). Disfonías funcionales (o disfuncionales). *Revista de logopedia, foniatría y audiología*, 2(1), 46-53.

Le Huche, F (1994). Exploración de la voz y del comportamiento fonatorio. En: Le Huche, F. & Allalí, A. (Editores). La voz. Tomo 2: *Patología vocal: semiología y disfonías disfuncionales* (pp. 5-19), París: Masson.

Le Huche, F., & Allali, A. (2003). La voz. Patología vocal de origen funcional, vol. 2. Barcelona: Masson.

Lieberman, P. (1968). Primate vocalizations and human linguistic ability. *The Journal of the Acoustical Society of America,* 44(6), 1574-1584.

Little, M. A., McSharry, P. E., Moroz, I. M., & Roberts, S. J. (2006). Testing the assumptions of linear prediction analysis in normal vowels. *J Acoust. Soc. Amer*., 119(1), 549-558.

Little, M. A. (2007). *Biomechanically informed nonlinear speech signal processing* (PhD thesis). Oxford University, UK.

Little, M. A., McSharry, P. E., Roberts, S. J., Costello, D. A., & Moroz, I. M. (2007). Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection. *Biomedical engineering online*, 6(1), 23.

Little, M. A., McSharry, P. E., Hunter, E. J., Spielman, J., & Ramig, L. O. (2009). Suitability of dysphonia measurements for telemonitoring of Parkinson's disease*. IEEE transactions on bio-medical engineering,* 56(4), 1015-1022.

López, L. M., & Cervilla, F. F. (2016). Epidemiology of vocal pathology among primary teachers. *Journal for Educators, Teachers and Trainers*, 7(1).

Marcotullio, D., Magliulo, G., & Pezone, T. (2002). Reinke's edema and risk factors: clinical and histopathologic aspects. *American journal of otolaryngology*, 23(2), 81-84.

Martins, R. H. G., Santana, M. F., & Tavares, E. L. M. (2011). Vocal cysts: clinical, endoscopic, and surgical aspects. *Journal of Voice*, 25(1), 107-110.

Martins, R. H. G., Pereira, E. R. B. N., Hidalgo, C. B., & Tavares, E. L. M. (2014). Voice disorders in teachers. A review. *Journal of voice*, 28(6), 716-724.

Menezes, L. N., Behlau, M., Gama, A. C. & Teixeira, L. C. (2011). Voice disorder clinic in the speech therapy outpatient unit at" Hospital das Clínicas da Universidade Federal de Minas Gerais". *Ciencia & saude coletiva*, 16(7), 3119-3129.

Michaelis, D., Gramss, T., & Strube, H. W. (1997). Glottal-to-noise excitation ratio–a new measure for describing pathological voices. *Acta Acustica united with Acustica*, 83(4), 700-706.

Milutinović, Z., & Vasiljević, J. (1992). Contribution to the understanding of the etiology of vocal fold cysts: a functional and histologic study. *The Laryngoscope*, 102(5), 568-571.

Moore, J.P. (1971). *Organic voice disorders*. Englewood Cliffs: Prentice-Hall.

Morente, J. C., Torres, J. A., Jiménez, M. C., Maroto, D. P., Rodríguez, V. P., Gomariz, E. M., ... & Ramos, A. J. (2001). Estudio objetivo de la voz en población normal y en la disfonía por nódulos y pólipos vocales. *Acta Otorrinolaringológica Española*, 52(6), 476-482.

Morrison, M. y Rammage, L. (1996). *Tratamiento de los trastornos de la voz*. Barcelona: Masson.

Müller, J. (1831). Beiträge zur Anatomie und Naturgeschichte der Amphibien. *Zeitschrift für Physiologie,* 4, 190-275.

Niebudek-Bogusz, E., & Śliwińska-Kowalska, M. (2013). An overview of occupational voice disorders in Poland. *International journal of occupational medicine and environmental health,* 26(5), 659-669.

Núñez-Batalla, F., Corte-Santos, P., Señaris-González, B., Llorente-Pendás, J. L., Górriz-Gil, C., & Suárez-Nieto, C. (2007). Adaptación y validación del índice de incapacidad vocal (VHI-30) y su versión abreviada (VHI-10) al español. *Acta Otorrinolaringológica Española*, 58(9), 386-392.

Núñez, F., Fernández, S. (2013). Afecciones benignas de las cuerdas vocales: lesiones exudativas del espacio de Reinke y otras lesiones. En Cobeta, I., Núñez, F. & Fernández, S (Eds). *Patología de la voz* (pp.242-257), Barcelona: Marge Médica Books

Núñez & Moreno (2013). Abuso y mal uso vocal. Valoración de la incapacidad vocal. En Cobeta, F., Núñez, I. & Fernández, S (Eds). *Patología de la voz* (pp.489-499), Barcelona: Marge Médica Books.

Oates, J. (2009). Auditory-perceptual evaluation of disordered voice quality: pros, cons and future directions. *Folia phoniatrica et logopaedica*, 61(1), 49-56.

Organización Mundial de la Salud (2001). Clasificación Internacional del Funcionamiento, de la Discapacidad y de la Salud. Madrid: Ministerio de Trabajo y Asuntos Sociales

Orlikoff, R. F. (1990). Heartbeat-related fundamental frequency and amplitude variation in healthy young and elderly male voices. *Journal of Voice*, 4(4), 322-328.

Paniagua, M. S., Cañete, M., Pérez, C. J., Salazar, C., Campos Roca, Y., De la O, E., Moreno, A. (2017). *Análisis acústico de las características de la voz en pacientes con Edema de Reinke y pólipos antes y después de la fonocirugía*. En *Vertentes e Desafios da Segurança* (pp 91-98). Leiria (Portugal): ASVDS – Associação Vertentes e Desafios da Segurança.

Paniagua, M. S., Pérez, C. J., Calle-Alonso, F., & Salazar, C. (2018). An Acoustic-Signal-Based Preventive Program for University Lecturers' Vocal Health. *Journal of Voice* (in press).

Perelló, J. (1962). La théorie muco-ondulatoire de la phonation. *Annales d'Oto-Laryngologie,* 79 (9), 722-725.

Perelló, J., Ponces, J. & Tresserra, L. (1977). *Trastornos del habla* (pp.25-45). Barcelona: Ed. Científico Médico.

Perelló, J & Salvá-Miquel, J. A. (1980). Alteraciones de la voz. Barcelona: Científico-Médica.

Perez, C.A, Preciado, J. (2003). Vocal fold nodules. Risk factors in teachers. A case control study design. *Acta Otorrinolaringológica Española*, 54, 253–260.

Pérez, C.A., Preciado, J., Preciado, M. (2003). Multi-dimensional study of the voice in the teaching professionals of La Rioja. *An Otorrinolaringol Ibero Am*., 30(4), 357-69.

Pérez, C.J., Paniagua M.S., Madruga, M., De la O, E. (2019). *Using linear and non-linear acoustic parameters to assess vocal status after phonomicrosurgery in Reinke's Edema.* En 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Berlín.

Pérez de Urbina, A. & Pérez, A. (2006). Manuel García, profesor de canto e inventor del laringoscopio y precursor de la laringoscopia (1805-1906). *Rev Patol Respir*, 9(3), 130-140.

Pickhard, A., & Reiter, R. (2013). Benigne Stimmlippentumore. *Laryngo-Rhino-Otologie,* 92(05), 304-312.

Prater, R. J., & Swift, R. W. (1984). *Manual of voice therapy*. Little Brown and Company.

Preciado, J., Pérez, C., Calzada, M., & Preciado, P. (2005). Frecuencia y factores de riesgo de los trastornos de la voz en el personal docente de La Rioja. Estudio transversal de 527 docentes: cuestionario, examen de la función vocal, análisis acústico y vídeolaringoestroscopia. *Acta Otorrinolaringológica Española*, 56(4), 161-170.

Preciado-López, J., Pérez-Fernández, C., Calzada-Uriondo, M., & Preciado-Ruiz, P. (2008). Epidemiological study of voice disorders among teaching professionals of La Rioja, Spain. *Journal of voice*, 22(4), 489-508.

R Development Core Team. (2014) *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing

Ramírez, M., & Marfil, I. (2014). *Anatomía y fisiología básica de la fonación*. En Montilla Ibáñez, M. A. (Ed.). Manual de Otorrinolaringología (pp. 1-8). Internet Medical Publising.

Ranchal, A., & Vaquero, M. (2008). Protocolo para la vigilancia de la salud del profesorado con atención a la enfermedad profesional. *Medicina y Seguridad del Trabajo*, 54(211), 47-60.

Real Decreto 1299/2006, de 10 de noviembre, por el que se aprueba el cuadro de enfermedades profesionales en el sistema de la Seguridad Social y se establecen criterios para su notificación y registro. BOE, nº 302, 19 de diciembre de 2006, pp 16-16.

Real Decreto Legislativo 8/2015 de 30 de octubre por el que se aprueba el texto refundido de la Ley General de la Seguridad Social. BOE, nº 261, de 31 de octubre de 2015, pp 80-80.

Refaeilzadeh, P., Tang, L., & Liu, H. (2009). Cross-validation. En Tamer Özsu, M. & Ling, L (Eds). *Encyclopedia of database systems*. (pp. 532-538), New York: Springer

Reveron, R. R., & ROMERO, R. (2007). Andreas Vesalius (1514-1564). Fundador de la Anatomía Humana Moderna. *Int. J. Morphol,* 25(4), 847-850.

Rodríguez-Parra, M. J., Adrián, J. A., & 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., Murry, T. (2000). Nomenclature of voice disorders and vocal pathology. *Otolaryngol Clinic North America*, 33, 1035-45.

Rosen, C. A., Lee, A. S., Osborne, J., Zullo, T., & Murry, T. (2004). Development and validation of the voice handicap index-10. *The Laryngoscope*, 114(9), 1549-1556.

Roy, N., Bless, D.M. (2000). Heisey D. Personality and voice disorder: a multitrait multidisorder analysis. *Journal of Voice*, 14, 521-48.

Roy, N., Bless, D.M. (2000). Personality traits and psychological factors in voice pathology: a foundation for future research. *J Speech Lang Hear Res*., 43, 737-48.

Roy, N., Merrill, R. M., Thibeault, S., Parsa, R. A., Gray, S. D., & Smith, E. M. (2004). Prevalence of voice disorders in teachers and the general population. *Journal of Speech, Language, and Hearing Research* 47(2), 281-293.

Roy, N., Barkmeier-Kraemer, J., Eadie, T., Sivasankar, M. P., Mehta, D., Paul, D., & Hillman, R. (2013). Evidence-based clinical voice assessment: a systematic review. *American Journal of Speech-Language Pathology*, 22, 212-226.

Rufo, M. J., Martín, J., Pérez, C. J., & Paniagua, S. (2019). A Bayesian decision analysis approach to assess voice disorder risks by using acoustic features. *Biometrical Journal,* 61(3), 503-513.

Russell, A., Oates, J., & Greenwood, K. M. (1998). Prevalence of voice problems in teachers. *Journal of Voice*, 12(4), 467-479.

Santana MC, Goulart BN, Chiari BM. (2012) Trastornos de la voz en profesores: revisión crítica de la práctica de vigilancia de la salud del trabajador. *J Soc Bras Fonoaudiol*, 24, 288–295.

Sataloff, R. T., Rosen, D. C., Hawkshaw, M., Spiegel, J. R. (1997) The three ages of voice. The aging adult voice. *Journal of Voice*, 11-2.

Señaris, B., Núñez Batalla, F., Corte, C. & Suárez, C. (2006). Índice de incapacidad vocal: factores predictivos. *Acta Otorrinolaringológica Española*, 57(2), 101-108.

Shigemori, Y. (1977). Some tests related to the air usage during phonation. *Clin Invest Otol,* 23, 138-66.

Simberg S., Sala E. y Ronnemaa A. M. (2004). A comparison of the prevalence of vocal symptoms among teachers, students and others university students. *Journal of Voice*, 18, 363-368.

Simberg, S., Sala, E., Tuomainen, J., Sellman, J., & Rönnemaa, A. M. (2006). The effectiveness of group therapy for students with mild voice disorders: a controlled clinical trial. *Journal of voice*, 20(1), 97-109.

Södersten, M., Granqvist, S., Hammarberg, B., & Szabo, A. (2002). Vocal behavior and vocal loading factors for preschool teachers at work studied with binaural DAT recordings. *Journal of Voice*, 16(3), 356-371.

Spiegel, J. R., Sataloff, R. T., & Hawkshaw, M. J. (1991). Strobovideolaryngoscopy: results and clinical value. *Annals of Otology, Rhinology & Laryngology*, 100(9), 725-727.

Stemple, J. C. (1993). *Principles of voice therapy. Voice therapy*: clinical studies. St. Louis: Mosby Year Book.

Stemple, J. (1999). Objetive measures of voice production in normal subjecct following prolonged voice use. *Journal of Voice*, 9, 127-133.

Tavares, E. L., & Martins, R. H. (2007). Vocal evaluation in teachers with or without symptoms. *Journal of Voice*, 21(4), 407-414.

Teixeira, J. P., & Gonçalves, A. (2014). Accuracy of Jitter and Shimmer measurements. *Procedia Technology*, 16, 1190-1199.

Theodoridis, S., & Koutroumbas, K. (2009). Linear Classifiers: Support Vector Machines. En: Theodoridis, S., & Koutroumbas, K. (Eds). *Patter Recognition* (pp. 91-150). San Diego, California (United States of America): Elsevier & AP.

Titze, I. R., Jiang, J. J., & Hsiao, T. Y. (1993). Measurement of mucosal wave propagation and vertical phase difference in vocal fold vibration. *Annals of Otology, Rhinology & Laryngology*, 102(1), 58-63.

Titze, I. R., & Winholtz, W. S. (1993). Effect of microphone type and placement on voice perturbation measurements. *Journal of Speech, Language, and Hearing Research*, 36(6), 1177-1190.

Titze, I. R. (1994). *Principles of Voice Production*. Prentice-Hall: Englewood Cliffs, NJ.

Titze, I.R. (1995). *Workshop on Acoustic Voice Analysis*. Denver: National Center for Voice and Speech.

Torres, B. (2007). *Anatomía funcional de la voz. Medicina del Canto*.

Torres, B (2013). La voz y nuestro cuerpo: un análisis funcional. *Revista de Investigaciones en Técnica vocal*, 1, 40-58.

Tsanas, A. (2012). *Accurate telemonitoring of Parkinson's disease symptom severity using nonlinear speech signal processing and statistical machine learning* (Doctoral dissertation), Oxford University, UK.

Tsanas, A., Little, M. A., McSharry, P. E., Spielman, J., & Ramig, L. O. (2012). Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease. *IEEE transactions on biomedical engineering*, 59(5), 1264-1271.

Tsanas, A., & Gómez-Vilda, P. (2013). Novel robust decision support tool assisting early diagnosis of pathological voices using acoustic analysis of sustained vowels. En *Multidisciplinary Conf*. *Users of Voice, Speech Sing (JVHC 13)* (pp. 3-12).

Urrutikoetxea, A., Ispizua, A., & Matellanes, F. (1995). Vocal pathology in teachers: a videolaryngostroboscopic study in 1046 teachers. *Revue de laryngologie-otologie-rhinologie*, 116(4), 255-262.

Van den Berg, J. (1958). Myoelastic-aerodynamic theory of voice production. *Journal of speech and hearing research*, 1(3), 227-244.

Van Houtte, E., Van Lierde, K., D'haeseleer, E., & Claeys, S. (2010). The prevalence of laryngeal pathology in a treatment-seeking population with dysphonia. *The Laryngoscope*, 120(2), 306-312.

Van Thal, J. H. (1961). Dysphonia. *Speech Path Ther*, 4, 1.

Vargas, M. A, & Muñoz, M.P, (2017). Características Vocales de un Grupo Docente: Universidad Metropolitana de Barranquilla. *Areté,* 17(2), 45-52.

Vesalius, A., Dalton, J., & Hartenfels, G. (1964). *De humani corporis fabrica libri septem.* Рипол Классик.

Vilaseca, I., Huerta, P., & Fernández, A.M. (2004). La voz tras la cirugía láser en el carcinoma glótico. En Bernal, M., Blanch, J. L. & Vilaseca, I. (Eds.) *Cirugía con láser CO2 en la vía aerodigestiva superior* (pp.375-389). Barcelona: Almirall.

Vilaseca, I., Huerta, P., Blanch, J. L., Fernández, A. M., Jiménez, C. & Bernal-Sprekelsen, M. (2008). Voice quality after CO2 laser cordectomy-What can we really expect? *Journal for the Sciences and Specialties of the Head and Neck*, 30(1), 43-49.

Webb, A. L., Carding, P. N., Deary, I. J., MacKenzie, K., Steen, N., & Wilson, J. A. (2004). The reliability of three perceptual evaluation scales for dysphonia. European *Archives of Oto-Rhino-Laryngology and Head & Neck*, 261(8), 429-434.

Williams NR. (2003). Occupational groups at risk of voice disorders: a review of the literature. *Occup Med (Lond),* 53(7), 456-60

Woisard-Bassols V. (2000). Bilan clinique da la voix. Encyclopedie Medico Chirurgicale. *Oto-Rhino-Laryngologie*. 20-753 A-10. Consultado el 26 de febrero de 2017 desde https://www.em-consulte.com/traite/OR/…/oto-rhino-laryngologie

Yumoto, E., Gould, W. J., & Baer, T. (1982). Harmonics-to-noise ratio as an index of the degree of hoarseness. *J Acoust Soc Am*, 71(6), 1544-1550.

Zagoruiko, N. G., & Tambovtsev, Y. (1982). Aspects of human performance in an intensive speech task. *International Journal of Man-Machine Studies*, 16(2), 173-181.

Zeitels, S. M., Bunting, G. W., Hillman, R. E., & Vaughn, T. (1997). Reinke's edema: phonatory mechanisms and management strategies. *Annals of Otology, Rhinology & Laryngology,* 106(7), 533-543.

Zhang, J., & Small, M. (2006). Complex network from pseudoperiodic time series: Topology versus dynamics. *Physical review letters*, 96(23), 238701.

Zraick, R. I., Risner, B. Y., Smith-Olinde, L., Gregg, B. A., Johnson, F. L., & McWeeny, E. K. (2007). Patient versus partner perception of voice handicap. *Journal of voice*, 21(4), 485-494.

Zraick, R. I., Kempster, G. B., Connor, N. P., Thibeault, S., Klaben, B. K., Bursac, Z., ... & Glaze, L. E. (2011). Establishing validity of the consensus auditory-perceptual evaluation of voice (CAPE-V). *American Journal of Speech-Language Pathology*, 20, 14-22.