

Análisis automatizado de eventos electroglotográficos. Una prueba piloto.

María Eugenia Torres¹
Juliana Codino²
Cristina Jackson-Menaldi³
Adam Rubin³

¹Facultad de Ingeniería
E-mail: metorres@santafe-conicet.gov.ar
Universidad Nacional de Entre Ríos.
CONICET

²Facultad de Medicina
Universidad de Buenos Aires

³Department of Otolaryngology, School of Medicine,
Wayne State University, MI, USA.

⁴Department of Otolaryngology-HNS,
University of Michigan Medical Center, MI, USA.

142



RESUMEN

La vibración de los pliegues vocales puede ser analizada de manera no invasiva a través del registro de los cambios de impedancia que ocurren en la glotis, mediante la técnica denominada electroglotografía. La señal correspondiente, el electroglotograma (EGG), es utilizada tanto en investigación biomédica como en la consulta clínica y su morfología está relacionada con diferentes patrones vibratorios. En particular, se refiere en la literatura reciente a la presencia de cierto tipo de oscilaciones, denominadas rodillas standard, sin que existan estudios cualitativos hasta el presente de este tipo de eventos. El objetivo de este trabajo es proponer un método automático de identificación y clasificación de los mismos en la porción ascendente y descendente del EGG con posible utilidad clínica. A tal efecto se desarrolló un algoritmo computacional basado en las propiedades matemáticas de dichas ondas, que permite la detección y reporte de las rodillas en la fase de contacto del EGG. El estudio se llevó a cabo mediante un análisis retros-

pectivo de sujetos adultos evaluados en un centro otorrinolaringológico de EEUU, en el cual se registraron EGGs como parte del protocolo de admisión. Los resultados obtenidos por el método aquí propuesto fueron comparados y analizados con los reportes manuales y visuales realizados por cinco expertos. Los resultados indican un nivel de concordancia óptimo, superando al esperado por azar. Esto permitiría proveer a los profesionales del área médica de un reporte automatizado, cuantita-

tivo y cualitativo, de los eventos presentes a lo largo de una emisión sin necesidad de analizar manualmente toda la señal EGG. Este método brindaría al especialista una mayor información que permitiría ganar en profundidad diagnóstica al ser asociado a la evaluación vocal standard. En el campo de la investigación vocal, esta herramienta también podría tener usos específicos.

1. INTRODUCCIÓN

La vibración de los pliegues vocales puede ser analizada directa o indirectamente a través de diferentes métodos de evaluación otorrinolaringológicos. A partir del siglo XIX se comenzó a hacerlo con el espejo de Manuel García, y los avances tecnológicos permitieron que esta práctica evolucionara hacia la videostroboscopia utilizando endoscopios rígidos y flexibles usados en la actualidad. Esa tecnología se fue superando con la videostroboscopia digital y de alta velocidad [1], permitiendo observar con un alto nivel de detalle las distintas fases del ciclo vibratorio en voces sanas como así también en casos de patología vocal.

También es posible evaluar la vibración de los pliegues vocales de manera no invasiva, registrando los cambios de impedancia que ocurren en la glotis, a través de la electroglotografía.

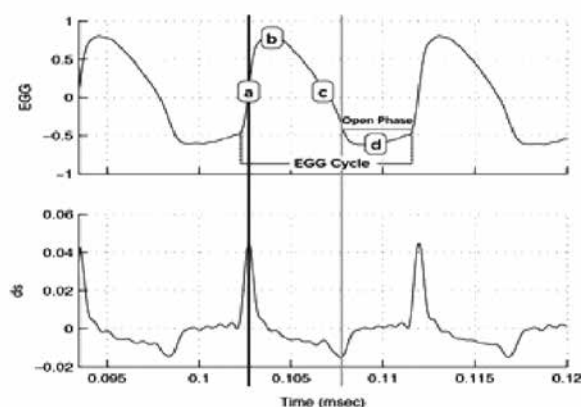


Figura 1: Ciclo EGG standard con las fases 'a' a 'd'. Uso de la derivada primera para determinar el comienzo (negro) y final (gris) de la fase de contacto.

La señal electroglotográfica (EGG) permite hacerlo teniendo en cuenta no sólo la imagen de la cara superior de los pliegues, sino además los cambios de la superficie medial y así caracterizar fácilmente aspectos relativos a las fases de contacto y pérdida de contacto de los mismos. Desde fines de los años '50 se la ha utilizado ampliamente tanto en investigación biomédica como en la consulta clínica considerando sus principales parámetros cuantitativos: determinación de los coeficientes de contacto [2, 3, 4, 5, 6, 7, 8, 9, 10] y de apertura [11, 12, 13], y como método de extracción de la frecuencia fundamental (F0) [6, 14, 15, 16]. Asimismo ha sido utilizada para obtener sincronización del disparo de la luz estroboscópica [17] y como elemento de retroalimentación en efectividad de la terapia vocal y/o médica [4, 5, 18].

A través del uso de modelos matemáticos de vibración de los pliegues vocales, simulando distintos patrones vibratorios, Titze logró reproducir las variaciones morfológicas de la señal [19]. Desde entonces se han utilizado cuatro variaciones morfológicas de señal EGG durante la evaluación clínica: peak skewing, skirt ramping, skirt bulging y pulse widening¹, representando cada una de ellas una desviación del patrón de normalidad.

En el EGG, un aumento en el contacto de los pliegues vocales se asocia a un incremento en la amplitud de la señal. La pendiente de la curva dependerá de la velocidad de contacto o pérdida de contacto de los pliegues vocales. Habitualmente la señal EGG presenta una pendiente más pronunciada en la fase de contacto (porción ascendente).

El ciclo electroglotográfico standard ha sido caracterizado por numerosos autores a lo largo de los años [12, 19, 20, 21, 22, 23, 24] (ver Figura 1):

Las fases descritas por estos autores son:

(a) Una porción ascendente con mayor pendiente, debido a una rápida fase de contacto.

¹ Dado que la clasificación hecha por el autor es en inglés, se mantendrá esta denominación en el idioma original.

(b) Un sector alrededor del pico máximo de la señal que representa la porción de máximo contacto entre los pliegues vocales.

(c) Una sección descendente con menor pendiente, que representa una pérdida de contacto gradual de los pliegues vocales. Este segmento puede o no presentar una "rodilla" y, de acuerdo a Titze [19] esta rodilla se relaciona con una aceleración en la separación de los pliegues vocales.

Este tipo de eventos serán llamados rodilla standard en este estudio.

(d) Una fase de apertura de los pliegues vocales representada por una porción de la señal con numerosas oscilaciones de baja amplitud. Esta fase indicará el final del ciclo glotal.

En la gran mayoría de las publicaciones se hace referencia a la mencionada rodilla standard (Figura 2 (a)). Sin embargo, es preciso aclarar que estas rodillas no se encuentran en todas las señales EGG (Figura 2 (b)). A través de diferentes tipos de análisis de señales EGGs en sujetos de ambos sexos, varios autores [2, 16, 25, 26, 27] encontraron presente la rodilla standard en mayor medida en hombres que en mujeres; y, al igual que otros estudios, describen una fuerte relación entre ese tipo de rodilla cóncava hacia abajo con el mínimo de la derivada primera (dEGG). Pero en muchas otras señales EGG, la morfología no se corresponde con la antes descrita: algunas presentan eventos (rodillas) en la porción ascendente (RodAsc) (Figura 2 (c)) o descendente (RodDesc) (Figura 2 (d)) de la señal. Sin embargo, de nuestro conocimiento, a la fecha aún no existen estudios cualitativos ni cuantitativos de este tipo de eventos de la señal EGG. Dado que la señal EGG da cuenta de los cambios de impedancia en la glotis, una más fácil identificación y descripción de los mismos permitiría revelar información oculta sobre la mecánica vibratoria de los pliegues vocales.

Teniendo en cuenta el estado del arte mencionado, en este estudio piloto retrospectivo, nuestro objetivo fue proponer un sistema automático de reconocimiento de rodillas en

la porción ascendente y descendente de la señal EGG. El presente estudio fue dividido en dos etapas:

- Etapa A: Desarrollar un método automático (MA) que reporte los eventos clínicamente relevantes en la porción ascendente y descendente de la señal EGG.

- Etapa B: Validar el desempeño del MA, estudiando el grado de acuerdo entre los resultados que se obtuvieron con el método y los reportados por el análisis visual realizado por expertos.

Se trabajó bajo la hipótesis de que el MA y el análisis visual realizado por un experto tienen un alto nivel de acuerdo, superior al esperado por azar.

Con este propósito, se confeccionó una base de datos de señales EGGs de acuerdo a lo descrito en la sección siguiente. Para validar el desempeño del MA, se estudió el nivel de acuerdo entre los resultados por él aportados y aquellos suministrados por los autores (CJM y JC) a través de inspección visual. Para estudiar la confiabilidad de las evaluaciones realizadas por los autores, tres licenciados en fonoaudiología externos al proyecto fueron convocados para actuar como evaluadores ciegos y realizar individualmente un análisis visual de los eventos que observaron en versiones impresas de la base de datos.

En el presente trabajo se incluyen y amplían resultados preliminares que fueron presentados en formato de poster en 2014 en The Voice Foundation 43rd Annual Symposium [28] (no publicados).

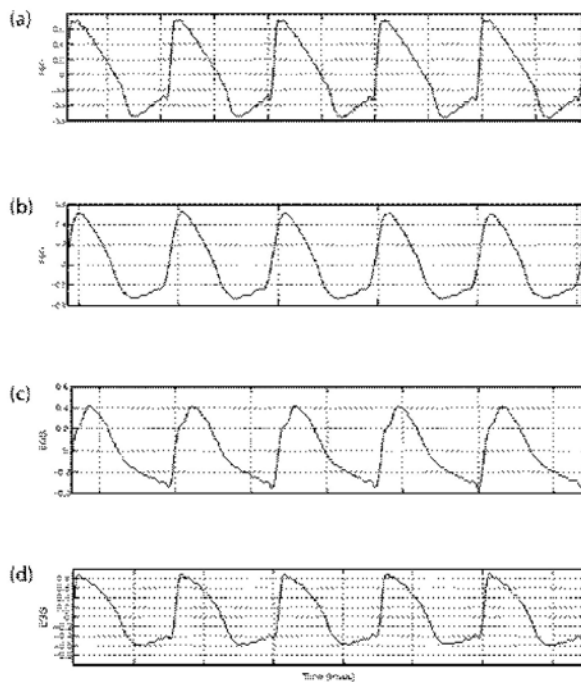


Figura 2: Señales EGG de sujetos de la base de datos del presente estudio comparando: (a) rodilla standard, (b) ausencia de rodilla standard, (c) RodAsc y (d) RodDesc

2. MATERIALES Y MÉTODOS

El presente trabajo se llevó a cabo mediante un estudio retrospectivo de sujetos adultos evaluados en Lakeshore Professional Voice Center (St. Claire Shores, MI. EEUU) por los autores CJM y AR. A cada sujeto se le asignó un número y esta fue su única identificación en la base de datos junto al diagnóstico laringológico, sexo y edad. Se contó con una base de datos correspondiente a señales EGGs obtenidas en el centro otorrinolaringológico mencionado. Todos los sujetos son adultos y fueron examinados por un médico ORL especialista en laringología (AR), y posteriormente derivados a una Dra. en fonoaudiología (CJM) quien llevó a cabo una evaluación vocal completa que incluye medidas EGGs. Se solicitó a cada sujeto que emita la vocal /a/ durante 2-4 segundos. Se seleccionó al azar un subgrupo de 310 pacientes de la base de datos y fueron analizados con el software Praat para Mac version 5.1.42, con el fin de identificar

y descartar aquellos con algún tipo de error de registro a causa de: desplazamiento de los electrodos durante la emisión; cuellos adiposos; uso excesivo de cremas o maquillaje en el día del estudio; o simplemente debido a voces severamente disfónicas. Luego de este análisis la primera base de datos quedó integrada por 212 señales EGGs. Todas las ondas EGGs seleccionadas fueron clasificadas según las cuatro morfologías de Titze [19]. Finalmente, una muestra de 25 señales EGGs de voces adultas, clasificadas como peak skewing y sin cambios morfológicos a lo largo de la grabación fueron seleccionadas al azar (ver Cuadro 1).

	Total
Edad	41,84 (8,8)
Mujeres	12
Hombres	13
Total ciclos EGC analizados	2951

Cuadro 1: Detalles de la muestra de señales EGGs utilizadas. *Promedio y entre paréntesis el desvío estándar.

La señal EGG fue registrada simultáneamente con la onda sonora utilizando el siguiente equipamiento: Laryngograph Kay Elemetrics, y la versión 7.0 del software Modelo 5138 Real Time EGG Analysis de dos canales. La señal acústica fue grabada con una frecuencia de muestreo de 44.100 Hz. La información de cada paciente fue codificada y los archivos [paciente_codigo].wav fueron incluidos en la muestra de este estudio.

2.1. Etapa A: Método Automático

Se desarrolló un algoritmo computacional y el correspondiente código fue implementado en Matlab© para detectar y extraer automáticamente eventos en la porción ascendente y/o descendente de la señal EGG. Este algoritmo fue diseñado con el objetivo de detectar los siguientes eventos en cada ciclo glotal (P):

- i. El comienzo (nCPhBeg) y el final (nCPhEnd) de la fase de contacto.

ii. La ubicación del tPmax del valor máximo de la señal (EGGPmax).

iii. La ocurrencia de RodAsc y RodDesc en la fase de contacto.

Esta información es automáticamente guardada e incluida en un archivo [output].xls. Para cada paciente, el correspondiente archivo contiene información relacionada con la muestra analizada, tiempo, amplitud EGG y tipo de evento detectado. En esta prueba piloto la fase de no contacto de los pliegues vocales fue excluida. El nCPhBeg fue determinado a través de la muestra en la cual se encontrara el máximo de la dEGG; y el nCPhEnd, a través del primer cruce por cero de la porción descendente del EGG.

En lo que sigue denominaremos a éste como el método Torres-Codino (MTC).

2.2. Etapa B: Validación

Se seleccionaron segmentos de 800ms de cada sujeto que fueron analizados por:

- El MTC (Figura 3)
- Análisis por inspección visual (“manual”)

por las autoras del presente estudio (CJM y JC). Los resultados de este análisis son considerados Gold Standard a los efectos del logro de los objetivos de la etapa correspondiente.

- Análisis por inspección visual (“manual”) ciega a los objetivos del trabajo, por tres evaluadores considerados expertos por tener experiencia en evaluación vocal y lectura de señales de voz.

3. RESULTADOS Y DISCUSIÓN

A los efectos del presente trabajo se compararon los resultados aportados por los expertos con los obtenidos automáticamente mediante MTC. Cabe señalar que el reporte manual consistió en identificar eventos sobre un total de 2951 períodos de las ondas de EGG correspondientes a las 25 señales de EGG, datos que posteriormente fueron volcados manualmente a sendas planillas electrónicas.

Para evaluar el acuerdo entre los métodos de análisis se calculó el coeficiente Kappa (k).

Koch y Landis [29] establecieron que el nivel de acuerdo puede ser caracterizado como pobre ($k = 0$ a 0.2), leve ($k = 0.21$ a 0.40), moderado ($k = 0.41$ a 0.60), considerable ($k = 0.61$ a 0.80), o casi perfecto ($k = 0.81$ a 1.0).

En el Cuadro 2, se observa la comparación de los resultados pareados para todos los métodos de análisis (MTC, gold standard, evaluador 1, evaluador 2 y evaluador 3). Ambos tipos de eventos muestran un alto nivel de acuerdo inter-evaluador, con valores de $k > 0.75$, confirmando la hipótesis de la etapa B (2.2) que el nivel de acuerdo entre el MTC y humanos es alto y superior al esperado por azar.

La fase de pérdida de contacto en la señal EGG presenta un mayor número de fluctuaciones que la fase de contacto. Varios autores describen las imprecisiones que conlleva obtener el instante de separación de los pliegues vocales basándose en la dEGG

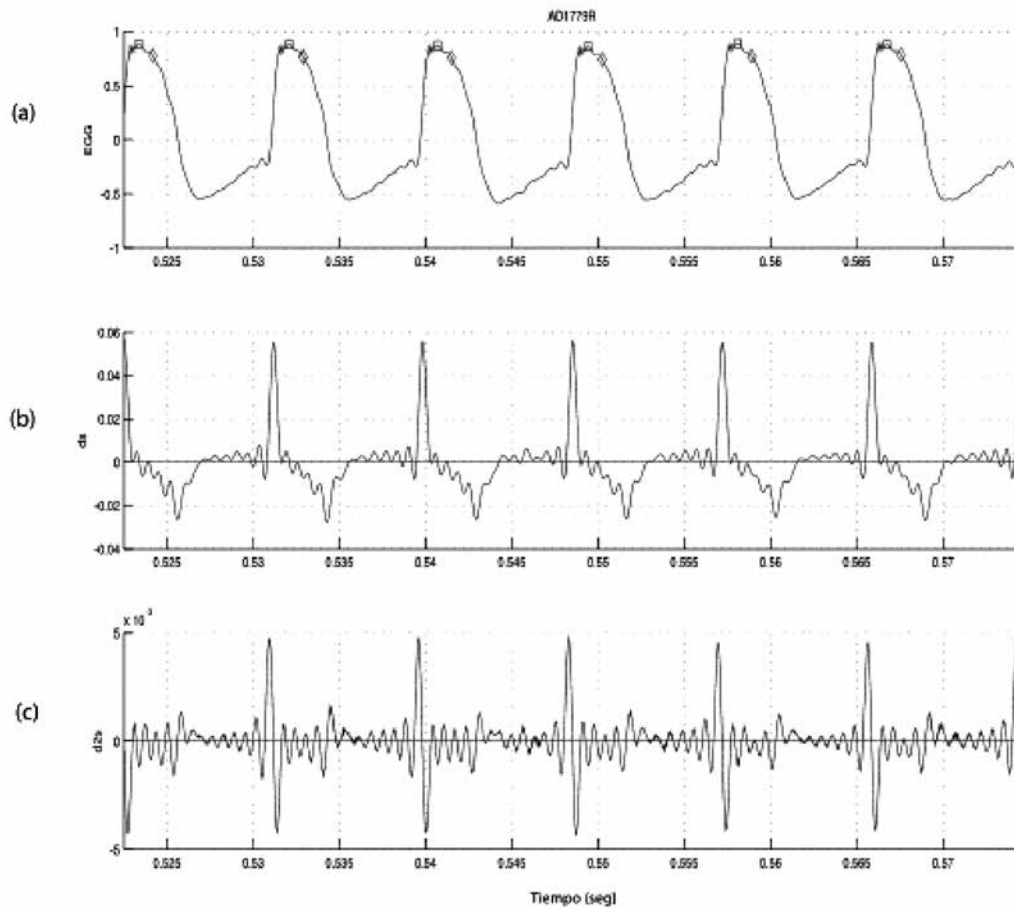


Figura 3: Metodo Torres-Codino. (a) EEG y puntos identificados en el reporte de MTC: máximo (cuadrado), estrella (RodAsc) y rombo (RodDesc). Primera (b) y segunda (c) derivadas del EGG.



[30, 31, 32, 12, 33]. Esto parece coincidir con el hecho de que la ubicación del mínimo de la dEGG es computacionalmente difícil debido al comportamiento oscilatorio de baja amplitud alrededor de los mínimos de la dEGG. En este estudio hemos superado esta dificultad calculando el nCPHend a través del primer cruce por cero de la porción descendente de la señal EGG en cada período.

Puede apreciarse en el Cuadro 2 que el grado de acuerdo es considerable, en el caso de la fase descendente, y casi perfecto en el de la fase ascendente. Además, en ambos casos el porcentaje de concordancia con las detecciones visuales realizadas por los expertos es alta, del orden del 90 %, por lo que puede concluirse que el enfoque utilizado es

adecuado y satisface los requisitos del objetivo propuesto.

	k no ponderado (EEM)	Concordancia observada (%)	Intervalo de confianza 95%
RodAsc	0,970 (0,0054)	98,67 %	0,9593; 0,9807
RodDesc	0,783 (0,0131)	89,27 %	0,7583; 0,8089

Cuadro 2: Valor estadístico de concordancia para los eventos RodAsc y RodDesc con el k no ponderado y el error estándar de la media (EEM) entre paréntesis, y su correspondiente IC (95 %).

4. CONCLUSIONES

Las señales EGGs de pacientes que presentan algún tipo de patología laríngea pueden clasificarse habitualmente en una de las cuatro morfologías propuestas por Titze [19]. Sin embargo, en algunos casos también exhiben determinadas rodillas en la porción ascendente y/o descendente de la señal, las cuales pueden contener información de valor clínico. En este trabajo hemos propuesto un algoritmo computacional que realiza automáticamente la detección y el reporte de este tipo de eventos a través del análisis computarizado de señales EGGs. El análisis estadístico realizado en una muestra de pacientes con diagnósticos laringológicos indica que el método propuesto es adecuado y confiable para reportar los eventos en las porciones de contacto ascendentes y descendentes.

Esta herramienta automatizada permitiría al clínico obtener un reporte cuantitativo y cualitativo de los eventos presentes a lo largo de una emisión sin necesidad de analizar manualmente toda la señal EGG.

Teniendo en cuenta que los resultados reportados en [34] corroboran los preliminares aquí presentados, como trabajos futuros se prevé extender esta prueba piloto a un mayor número de casos que cumplan con los criterios de inclusión y avanzar hacia el análisis de la posible relación de estos eventos con la información clínica de los pacientes, obrante en la base de datos.

5. REFERENCIAS

[1] Qin X, Wang S and Wan M 2009 IEEE Transactions on Bio-Medical Engineering 56 1744-1754 ISSN 1558-2531 URL <http://www.ncbi.nlm.nih.gov/pubmed/19272979>

[2] Awan S N and Awan J A 2013 Journal of Voice 27 433-440 ISSN 0892-1997 URL <http://www.sciencedirect.com/science/article/pii/S0892199713000532>

[3] Master S, Guzman M, Carlos de Miranda H and Lloyd A 2013 Journal of voice: official journal of the Voice Foundation 27 187-194 ISSN 1557-8658

[4] Ayazi S, Pearson J and Hashemi M 2012 Journal of clinical gastroenterology 46 119-123 ISSN 1539-2031

[5] Gaskill C S and Quinney D M 2011 Journal of Voice: Official Journal of the Voice Foundation ISSN 1557-8658 URL <http://www.ncbi.nlm.nih.gov/pubmed/21550779>

[6] Daffern H and Howard D M 2010 Logopedics, phoniatrics, vocology [http://www.jvoice.org/article/S0892-1997\(15\)00238-6/abstract](http://www.jvoice.org/article/S0892-1997(15)00238-6/abstract) 68-73 ISSN 1651-2022

[7] Ma E P M and Love A L 2010 Journal of Voice: Official Journal of the Voice Foundation 24 146-152 ISSN 1557-8658 URL <http://www.ncbi.nlm.nih.gov/pubmed/19481415>

[8] Zagolski O 2009 Journal of Voice: Official Journal of the Voice Foundation 23 567-571 ISSN 1557-8658 URL <http://www.ncbi.nlm.nih.gov/pubmed/18504110>

[9] Mayes R W, Jackson-Menaldi C, Dejonckere P H, Moyer C A and Rubin A D 2008 Journal of voice: official journal of the Voice Foundation 22 756-759 ISSN 1873-4588

[10] Herbst C and Ternstrom S 2006 Logopedics, Phoniatrics, Vocology 31 126-138 ISSN 1401-5439 URL <http://www.ncbi.nlm.nih.gov/pubmed/16966155>

[11] Echternach M, Dippold S, Sundberg J, Arndt S, Zander M F and Richter B 2010 Journal of voice: official journal of the Voice Foundation 24 644-650 ISSN 1873-4588

[12] Henrich N, d'Alessandro C, Doval B and Castellengo M 2004 The Journal of the Acoustical Society of America 115 1321-1332 ISSN 0001-4966

[13] Childers D G, Hicks D M, Moore G P, Eskenazi L and Lalwani A L 1990 Journal of speech and hearing research 33 245-254 ISSN 0022-4685

[14] Faria B S d, Oliveira K V d, Silva J P G e, Reis C, Ghio A and Gama A C 2012 Brazilian journal of otorhinolaryngology 78 29-34 ISSN 1808-8686

[15] Svec J G, Popolo P S and Titze I R 2003 Logopedics, phoniatrics, vocology 28 181-192 ISSN 1401-5439

[16] Chen Y, Robb M P and Gilbert H R 2002 J o Speech, Language, and Hearing

- Research: JSLHR, 45821-829 ISSN 1092-4388 URL: <http://www.ncbi.nlm.nih.gov/pubmed/12381041>
- [17] Dejonckere P H 2000 Otolaryngologic clinics of North America 33 731-750 ISSN 0030-6665
- [18] Herbst C T, Howard D and Schlomicher-Thier J 2010 Journal of Voice: 24 72-85 ISSN 1557-8658 URL <http://www.ncbi.nlm.nih.gov/pubmed/19185453>
- [19] Titze I R 1990 Journal of voice 4 1-9 URL <http://www.sciencedirect.com/science/article/pii/S0892199705800761>
- [20] Fourcin A 2010 Logopedics, Phoniatrics, Vocology 35 74-80 ISSN 1651-2022 URL <http://www.ncbi.nlm.nih.gov/pubmed/20536379>
- [21] Colton R H and Conture E G 1990 Journal of Voice 4 10-24 ISSN 0892-1997 URL [http://www.jvoice.org/article/S0892-1997\(05\)80077-3/abstract](http://www.jvoice.org/article/S0892-1997(05)80077-3/abstract)
- [22] Robert Orlikoff 1998 Phonoscope 1
- [23] Childers D G and Larar J N 1984 IEEE transactions on bio-medical engineering 31 807-817 ISSN 0018-9294
- [24] Herbst C T, Fitch W T S and Svec J G 2010 The Journal of the Acoustical Society of America 128 3070-3078 ISSN 1520-8524 URL <http://www.ncbi.nlm.nih.gov/pubmed/21110602>
- [25] Awan S N, Krauss A R and Herbst C T 2015 Journal of Voice ISSN 0892-1997 URL <http://www.sciencedirect.com/science/article/pii/S089219971400246X>
- [26] Titze I R 1989 Journal of the Acoustical Society of America 85 1699-1707
- [27] Robb M P and Simmons J O 1990 Journal of the Acoustical Society of America 88 1318-1322
- [28] Codino J, Torres M E, Moore R, Rubin A and Jackson Menaldi C 2014 Automated electroglottographic shape analysis: A pilot study The Voice Foundation. 43rd Annual Symposium Care of the Professional Voice (unpublished)
- [29] Koch G and Landis J 1977 Biometrics 33 159-174
- [30] Baer T, Lofqvist A and McGarr N S 1983 The Journal of the Acoustical Society of America 73 1304-1308 ISSN 0001-4966
- [31] Titze I R and Scherer R C 1983 Vocal Fold Physiology. Biomechanics, Acoustic and Phonatory Control (The Denver Center for the Performing Arts)
- [32] Moore D M, Berke G S, Hanson D G and Ward P H 1987 Laryngoscope 97 543-553
- [33] Herbst C T, Lohscheller J, Svec J G, Henrich N, Weissengruber G and Fitch W T 2014 J Exp Biol 217 955-963 URL <http://dx.doi.org/10.1242/jeb.093203>
- [34] Codino J, Torres M, Rubin A and Jackson-Menaldi C 2016 Journal of Voice: Official Journal of the Voice Foundation (In Press) URL [http://www.jvoice.org/article/S0892-1997\(15\)00238-6/](http://www.jvoice.org/article/S0892-1997(15)00238-6/)