

MÁSTER INTERUNIVERSITARIO POR LA UNIVERSIDAD AUTÓNOMA DE
BARCELONA Y LA UNIVERSIDAD RAMON LLULL

Coordinadores

Carrasumada Serrano Pau y Josep Maria Vila Rovira

**Trabajos preliminares sobre la validación del
programa de análisis acústico Vocalab 4 como
instrumento de evaluación de voces aperiódicas**

MÁSTER EN TRASTORNOS DE LA COMUNICACIÓN Y DEL LENGUAJE

-FACULTAD DE PSICOLOGÍA -

Trabajo Final de Máster por **Felipe Pulla Fajardo**

septiembre del 2016

TUTOR: **Dr. Josep Maria Vila-Rovira**

ÍNDICE

1. Resumen.....	Pág. 3
2. Introducción.....	Pág. 4
2.1 Calidad de vida.....	Pág. 4
2.2 Calidad vocal.....	Pág. 5
2.3 Mecanismos sustitutorios de habla.....	Pág. 5
2.4 Análisis perceptivo.....	Pág. 7
2.5 Análisis acústico.....	Pág. 7
3. Método.....	Pág. 9
3.1 Participantes.....	Pág. 9
3.2 Material y procedimientos.....	Pág. 9
4. Resultados.....	Pág. 11
5. Discusión y conclusiones.....	Pág. 14
6. Referencias bibliográficas.....	Pág. 16

1. Resumen

Las voces de pacientes con laringectomía total presentan dificultad en su evaluación. Estas evaluaciones se realizan a través de tests perceptivos y de programas de análisis acústico, cuyos valores obtenidos aportan cierta relevancia para el profesional clínico de la voz. Esta investigación pretende identificar la validez de los estudios objetivos en las voces esofágicas y traqueoesofágicas. Se ha estudiado la voz de 31 personas con laringectomía total con y sin prótesis fonatoria. Los resultados obtenidos muestran diferencias significativas en diversos parámetros perceptivos en relación a los parámetros acústicos con una significación $p < .05$. Cuestionando la efectividad de los programas de análisis acústico como herramienta evaluadora pero a su vez ser capaces de registrar valores altamente relacionados con el oído humano.

Palabras clave: Laringectomía total; Voz esofágica; Análisis perceptivo; Análisis acústico

Resum

Les veus de pacients amb laringectomia total presenten dificultat en la seva avaluació. Aquestes avaluacions es realitzen a través de tests perceptius i de programes d'anàlisi acústica, els valors obtinguts aporten certa rellevància pel professional clínic de la veu. Aquesta investigació pretén identificar la validesa dels estudis objectius en les veus esofàgiques i traqueoesofàgiques. S'ha estudiat la veu de 31 persones amb laringectomia total amb i sense pròtesi fonatòria. Els resultats obtinguts mostren diferències significatives en diversos paràmetres perceptius en relació als paràmetres acústics amb una significació $p < .05$. Qüestionant l'efectivitat dels programes d'anàlisi acústica com a eina avaluadora però al seu torn ser capaços de registrar valors altament relacionats amb l'oïda humana.

Paraules clau: Laringectomia total; Veu esofàgica; Anàlisi perceptiu; Anàlisi acústica

Abstract

The voices of patients with total laryngectomy have difficulty in evaluation. These evaluations are done through tests perceptive and acoustic analysis programs, the values obtained provide some relevance to clinical professional voice. This research aims to identify the validity of objective studies in esophageal and tracheoesophageal voices. We have studied the voice of 31 people with total laryngectomy with and without voice prosthesis. The results show significant differences in various parameters in relation to perceptive acoustic parameters with a significance $p < .05$. Questioning the effectiveness of acoustic analysis programs as evaluative tool but in turn be able to record values highly related to the human ear.

Keywords: Total laryngectomy; esophageal voice; perceptual analysis; acoustic analysis

2. Introducción

El procedimiento quirúrgico que incluye la extirpación completa de la laringe, tras un cáncer avanzado, es la laringectomía total (LT). En esta intervención se produce la extracción de las cuerdas vocales, las cuales son las encargadas en producir los sonidos del habla. Obteniendo como resultado, un serio deterioro en la comunicación. (Ferrat & Guerti, 2012). Además de la pérdida de la voz, puede haber pérdida de olfato, un pobre reflejo de la tos, dificultad para tragar, cambios pulmonares y las complicaciones asociadas con un traqueostoma permanente (orificio por el que la tráquea se comunica con el exterior).

Entre las modificaciones anatómicas, fisiológicas y funcionales mencionadas, se destaca en especial el impacto que provoca la pérdida total de la voz, la ausencia de este sistema de comunicación hace dimensionar aún más su riqueza funcional. Siendo una herramienta importante de la vida, del día a día, en un mundo que se basa, en gran medida, en la comunicación oral. Por este motivo, la rehabilitación funcional de los pacientes laringectomizados sigue siendo una de las áreas de mayor interés para cirujanos de cabeza y cuello, además de suponer un gran reto para los logopedas en la instauración de un nuevo mecanismo de habla.

2.1 Calidad de vida

Cabe destacar que dicha intervención provoca además una gran merma en relación a la calidad de vida (CV). Sí es verdad que en las tres últimas décadas se observan varios avances en la rehabilitación del habla dando lugar a mejoras en la CV de estos pacientes. (Deore, N., Datta, S., Dwivedi, R., Palav, R., Shah, R., Sayed, S., Jagde, M. & Kazi, R., 2011). Varios estudios como el de Fomby & Cherlin (2011) y Doyle (2015) aseguran que dicha la CV está generalmente afectada durante los primeros meses después de la LT, pero al cabo de doce meses, muchos pacientes se adaptan a su nueva situación. Si bien esta es la parte positiva de los resultados, también hay otra parte de esta población que experimenta, a largo plazo, las consecuencias de esta falta de adaptación (Hassan & Weymuller, 1993), lo que informa sobre las dificultades añadidas que dichos pacientes sufren. Como resultado de ello, es importante identificar los factores que predicen la CV de modo que los objetivos puedan ser identificados para proceder con la rehabilitación. Contando con predictores tradicionales, como la edad, sexo, tipo de discurso sin laringe

o la enfermedad del paciente pero sin olvidarse también de contemplar variables psicosociales y el cómo afrontan esta nueva situación.

2.2 *Calidad vocal*

En cuanto a la calidad de la voz se debe saber que es un componente multidimensional, el cual aporta mayor fluidez al discurso además de obtener una amplia aplicación tanto cultural como fisiológica, jugando un papel esencial en la comunicación oral. Por lo que dicha calidad, puede ayudar o dificultar la inteligibilidad, siendo también una fuente rica de información indicativa sobre determinantes lingüísticos, culturales y familiares. Por tanto, si se consigue un rápido restablecimiento de una voz aceptable y fluida con habla inteligible, en pacientes laringectomizados, se obtendrá un proceso fundamental para el éxito de la rehabilitación y adaptación psicológica en estos pacientes (Deore et al., 2011).

Tal éxito se consigue al combinar una óptima intervención logopédica en conjunto con evaluaciones sistemáticas, tanto del tipo perceptivas como objetivas durante todo el proceso rehabilitador. Siendo estas las encargadas de aportar *feedback* sobre el tipo de calidad que se va obteniendo.

2.3 *Mecanismos sustitutorios de habla*

Para llevar a cabo dicha intervención logopédica, se debe conocer qué tipos de mecanismos sustitutorios de habla son los empleados en pacientes laringectomizados. Existen tres principales herramientas de vocalización, los cuales darán pie a la instauración de un nuevo mecanismo vocal después de una LT. Dichos mecanismos son: la prótesis de voz traqueoesofágica, el habla esofágica y la electrolaringe. En primer lugar cada paciente debe conocer el tipo de mecanismo, al que es candidato. Para ello, hay que tener en cuenta la características individuales de cada uno, principalmente las relacionadas con la edad y condiciones físicas, el origen social, el tipo de cirugía (por ejemplo, en combinación con glosectomía y/o disección del cuello) o la radioterapia (Tóth, Csernoch, Sziklai, & Szűcs, 2014).

En relación a los diferentes tipos de mecanismos, la voz esofágica (ES) es la más utilizada en nuestro país. La cual consiste en aprender a producir el sonido en la entrada del esófago, a través de diferentes métodos rehabilitadores (Mart, Barandiar, & Fern, 2006). Dicho sonido resultante es parecido al del eructo, como el que se hace involuntariamente tras ingerir alimentos o bebidas gaseosas. En la entrada del esófago se encuentra una

pequeña membrana que puede vibrar produciendo sonido si un soplo de aire lo atraviesa al subir. Este será su vibrador de reemplazo, que recibe el nombre de “neoglottis”, nueva glottis o pseudoglottis. Tras la operación ya no podrá utilizar la reserva de aire de los pulmones para proyectar aire hasta la boca y emitir sonidos. A partir de ahora su esófago se encargará de hacer reservas de aire, como si fuera un globo alargado. De ahí que esta forma de hablar sea conocida como “voz esofágica”. A veces se refiere a ella más técnicamente con el nombre de “voz erigmofónica”.

Por otro lado, se encuentra la voz traqueoesofágica (TE). Este tipo de voz es más fácil de adquirir que la ES, pero sólo es un recurso posible para los operados a quienes han practicado una fístula traqueoesofágica o aquellos que han colocado un implante fonatorio (o prótesis fonatoria interna). El sonido se produce igualmente en la entrada del esófago, pero es más fácil de producir, modular y prolongar, ya que es alimentado por el aire de los pulmones. Dicho paciente debe obturar el orificio traqueal, desviando hacia el esófago el soplo del aire, por medio de un pequeño tubo situado entre éste y la tráquea; por esta razón llamamos a esta voz traqueoesofágica.

Y la última opción es la electrolaringe, una prótesis externa cuyo funcionamiento es parecido al de un birimbao, un pequeño instrumento de música cuyo sonido es generado por una lámina vibratoria, modulado en la boca en función de la posición de la lengua. Incluye un generador de sonido que reemplaza la voz. Por lo que este sonido es conducido hasta la cavidad bucal, donde se transforma en habla y se hace inteligible cuando el paciente susurra. Es un sonido producido gracias a la eléctrica, a través de una batería de energía que pone en marcha la actividad de su vibración. (Heuillet-Martin, 2003).

La principal diferencia en las distintas técnicas de vocalización, a excepción de la electrolaringe, difiere en la forma en que se almacena el aire. Para la fonación TE, el aire derivado de los pulmones aumenta la presión por debajo del segmento faringoesofágico (FE). Cuando la presión del segmento sub-FE es suficientemente alta, la neoglottis se pone en vibración. En el caso de la fonación ES, la parte superior del esófago sirve como el nuevo depósito, y el segmento FE vibra cuando una presión intra-ES es suficientemente positiva. Debido a la complejidad del segmento FE, y la falta de control completo sobre la nueva estructura vibratoria, las voces ES y TE parecen ser ásperas, roncadas y débiles, al percibirlo por los oyentes. (Yan, N., Ng, M. L., Wang, D., Zhang, L., Chan, V., & Ho, R. S., 2013). Aunque esas características acústicas, a efectos de calidad vocal, aumentan en las voces ES. Debido a que las grandes superficies de vibración de la mucosa y el tracto

vocal, mayormente alterado, debilitan la función resonadora, por lo que el tono del habla esofágica es bajo, generalmente entre 50 y 100 Hz. Además, inicialmente la alteración de tono de voz y los aspectos prósodicos son muy difíciles de controlar (Tóth, A., Csernoch, L., Sziklai, I., & Szűcs, A., 2014).

2.4 Análisis perceptivo

Para realizar la valoración perceptiva de la calidad vocal en esa nueva voz, se dispone de un test de evaluación perceptiva “INFVo” (Moerman, 2006), test creado expresamente para personas que han padecido un proceso de extirpación de laringe y han tenido que instaurar un nuevo sistema de comunicación oral. Moerman diseñó una escala de valoración perceptiva que responde a las siglas INFVo, donde la I (Intelligibility) supone la cantidad de esfuerzo que sería necesario para entender una vocal /a/ sostenida, la N (Noise) evalúa la cantidad de molestia causada por la audibilidad del estoma, la letra F (Fluency) refleja la estabilidad percibida en la producción del sonido /a/ y finalmente, la Vo (Voicing) valora la calidad de la voz esofágica o traqueoesofágica.

El experto puntúa el nivel de severidad de los rasgos vocales, de esta nueva herramienta, en una escala de cuatro puntos. Donde al primero se le asigna el valor 0 que responde al máximo valor negativo de los parámetros y los valores 1, 2 y 3 indican los niveles de mejora, cuanto mayor sea la puntuación mejor será la calidad de cada uno de los parámetros valorados. Dicha escala perceptiva no es muy conocida en nuestro país, por lo que los datos que se conocen, sobre su uso, son los aportados por estudios realizados en Alemania, país de origen del autor del test

2.5 Análisis acústico

Otra de las evaluaciones a realizar son las objetivas. Para ello, se utiliza el análisis acústico a través de programas de software. Hace alrededor de unas cuatro décadas que se incorporó este tipo de análisis, (Maryn, Roy, De Bodt, Van Cauwenberge, & Corthals, 2009), con el fin de objetivar la evaluación en un fenómeno, tan poco tangible, como es la voz. El uso de este tipo de análisis es comúnmente aplicado en ámbitos clínicos, ya que dichos softwares, encargados de la evaluación, aportan a los parámetros de estudio valores numéricos. De ahí que se reconozca como una prueba de evaluación objetiva, garantizando así su imparcialidad en el análisis, abstrayéndose de consideraciones subjetivas. Aparentemente parecía ser un buen instrumento de trabajo pero actualmente aumenta su controversia, aún así se sigue dando uso como complemento de la evaluación

y como *feedback* de los programas de intervención aplicados. Por lo que gracias a los avances de la informática, se consigue realizar un análisis de la voz, facilitando a los clínicos e investigadores el uso de soportes que aportan datos sobre el sonido de la voz. Siendo este tipo de análisis, simplemente, otra visión del estudio vocal.

A través de estos programas informáticos específicos, como puede ser Vocalab 4, Praat, MDVP, Doctor Speech, se encuentran una serie de parámetros que varían en tipo y número, dependiendo de qué programa se esté utilizando y de la profundidad del análisis deseado. Los más estudiados y con mayor recorrido en la clínica habitual son: la frecuencia fundamental (F_0), tiempo máximo de fonación (TMF), referentes a la perturbación de frecuencia (Jitter), este es uno de los más cuestionados hasta ahora junto con los referentes a la perturbación de amplitud (Shimmer), y los referentes a los armónico-ruido (HNR). Dichos parámetros aportan datos que son comparados con valores estándar, previamente baremados, de voz laríngea. Por lo que a través del análisis acústico de voces aperiódicas, tipo III (Titze, 1995), se podrá conocer cuánto distan los datos obtenidos, en dichas voces, con respecto a los valores estándar de voz laríngea. Tras los resultados se puede inferir, que impacto tendrá esta nueva voz y si con ella conseguimos reducir aquellas dificultades encontradas posteriormente a la cirugía.

Aunque también el análisis acústico informatizado presenta limitaciones al intentar analizar señales acústicas altamente aperiódicas. Como es el caso de las voces esofágicas o traqueoesofágicas. Aunque la nueva versión del programa Vocalab 4, remite haber realizado un ajuste de algoritmos con el que pretende mejorar los registros de los parámetros en este tipo de voces que presentan tanta dificultad de análisis.

Por consiguiente y una vez conocidos los datos que se aportan, se plantea como objetivo en la investigación conocer si el programa Vocalab 4, es un buena herramienta de evaluación para voces aperiódicas. Planteando como hipótesis, conseguir que alguno de sus parámetros registre valores que coincidan, de forma próxima, con los aportados por el oído humano.

3. Método

3.1 Participantes

Para este estudio, la muestra experimental se seleccionó entre pacientes laringectomizados totales con y sin prótesis fonatoria. Los cuales fueron intervenidos por las diferentes unidades de cáncer de cabeza y cuello, de los servicios de ORL, en los diferentes hospitales distribuidos por todo el territorio catalán. Estas unidades, una vez que seleccionan al paciente para ser laringectomizados, activan un protocolo de actuación. Dicho protocolo consiste en comunicar a la asociación de laringectomizados, correspondiente según la provincia, el nuevo caso de LT. El objetivo de este hecho es que un representante laringectomizado de la asociación, con habla esofágica, acuda al hospital una vez se haya producido la intervención quirúrgica. Con el fin de presentar un modelo postquirúrgico, y así apoyar en esos momentos de desconcierto para el paciente. También podrá conocer cuáles son los pasos a seguir y acompañarlo en el proceso rehabilitador.

Tras el criterio de selección de la muestra experimental se obtuvieron un total de 31 pacientes laringectomizados totales. El conjunto estaba formado por 30 hombres, llevando 3 de ellos prótesis traqueoesofágica, y 1 mujer, con una media de edad del total de la muestra de 68 años.

La participación al estudio fue totalmente voluntaria, previo cumplimentación y firma de los permisos pertinentes para la correcta realización del mismo. Se administraron pruebas que son de uso habitual en nuestra práctica clínica y de las que no existe evidencia alguna acerca de su invasividad ni sobre su nocividad. Para asegurar la privacidad de los pacientes, se les atribuyó un código alfanumérico que garantizaba el anonimato de los participantes al estudio.

3.2 Materiales y procedimientos

La recogida de muestras se realizaron durante los meses de marzo a mayo, a través de sesiones individuales con una media, aproximada, de 10 minutos por paciente. Se utilizó una sala no insonorizada pero con un nivel de sonoridad inferior a los 50 dB. En ella, se procedió a la obtención de datos de filiación y a los registros relevantes para el estudio. Dicha recogida se centró en el registro de la vocal /a/ de manera sostenida con la mayor

extensión posible, obteniendo tres vocales por pacientes haciendo un total de 93 grabaciones.

La obtención de registro de las vocales se realizó mediante un micrófono de diadema AKG 555, un grabador digital Audacity, the Free, Cross-Platform Sound Editor y una tarjeta de sonido, Scarlett 2i4 que garantizara la calidad de las grabaciones. Para conocer el nivel de sonoridad de la sala empleada para los registros se utilizó un sonómetro digital PCE-MSM 2. El formato digital elegido para las grabaciones fue *wave* con una frecuencia de 44.100 Hz a 16 bits y en mono. Las muestras fueron analizadas por el programa de análisis acústico, VOCALAB 4. Se hizo un primer análisis con el fin de seleccionar qué audio, de los tres obtenidos por cada paciente, marcaba mejores resultados en los parámetros de estudio y así conseguir nuestra muestra final, la cual se compuso por 31 audios. El programa empleado cuenta con 20 años de trayectoria dentro del mundo del análisis acústico pero desconocido en la práctica habitual de la logopedia española. Anne y Etienne Sicard, naturales de Francia, logopeda e ingeniero informático respectivamente, son los creadores de dicho soporte. Con él han mejorado desde la parte más pedagógica de su uso en dichos software, hasta el ajuste de algoritmos implicados en el análisis de la voz patológica. Con el fin de obtener un mejor manejo con alta carga intuitiva y mayor precisión en los resultados de análisis. Indicando mayor capacidad para obtener mejores resultados en relación a las voces aperiódicas, en comparación a otros programas de análisis acústico (Menin-Sicard & Sicard, 2016) .Los parámetros obtenidos por este programa fueron alteración de ataque de sonido, inestabilidad de altura (Jitter), inestabilidad de amplitud (Shimmer), ruido/señal y pobreza armónica (HNR).

Esta investigación cuenta con un jurado de expertos en el análisis subjetivo de voces esofágicas, el cual está compuesto por 5 logopedas especialistas en alteraciones de la voz. Dicho jurado hizo uso del test de evaluación subjetiva “INFVo” de Moerman (2006), el cual fue diseñado para evaluar voces de pacientes que han sufrido la sustitución de su voz, como es el caso de la muestra escogida. Deberán emplear una escala de 0 a 3 y puntuar los parámetros vocales; Inteligibilidad, Ruido, Fluidez y Sonoridad. Teniendo en cuenta la definición atribuida por los autores del test con respecto a los parámetros.

Los datos analizados se han tratado con el paquete estadístico IBM SPSS, versión 22, con él se han calculado la consistencia interna, a través del análisis de fiabilidad α Cronbach para el test perceptivo “INFVo” cuyo fin era conocer la covarianza entre sus puntuaciones y observar si el conjunto del jurado media de igual forma, los diferentes parámetros.

Además se realizaron estadísticos descriptivos del tipo media y desviación estándar para ambas herramientas de evaluación. Y a su vez, se realizó también un análisis de correlación de Pearson entre las variables perceptivas y acústicas con el fin de determinar si existe algún tipo de significación entre ambas $p < .05$.

4. Resultados

Los resultados obtenidos con nuestra población de estudio, cuyas características se encuentran recogidas en la tabla 1, se presentaran organizados por diferentes bloques.

Tabla 1. Características de la población estudiada

Varibles	%
Género	
Hombre	78.9
Mujeres	2.6
Edad	68.32*
Tratamiento	
Ninguno	15.8
Quimioterapia	18.4
Radioterapia	26.3
Quimio y Radio	21.1
Uso de prótesis	
No	73.7
Sí	7.9
Año intervención	2013**

Datos según frecuencia. *Promedio de edad.

**Moda sobre el año de intervención

En estos resultados se hace referencia a los estadísticos descriptivos resultantes del test “INFVo”, en el que sus medias y desviaciones estándar nos indican la poca variabilidad que existe entre las puntuaciones de los parámetros perceptivos, registrado por los jueces expertos (datos a observar en la tabla 2). A raíz de estos datos, se permite someter sus valores al análisis de fiabilidad de α Cronbach, con el que se obtiene un resultado de $\alpha = .944$. Afirmando así, que todos los jueces expertos median de igual forma los diferentes parámetros del test. Por tanto, esta alta consistencia interna posibilitó la creación de la variable “Media” para cada uno de sus parámetros, cuyo valor fue el promedio de sus puntuaciones. Además se creó también otra variable que permitió obtener un único valor total del test, a la que se llamó “Valor global perceptivo” (VGP). Dicha nueva variable se decidió crear al conocer también, que el grado de significación entre las variables perceptivas era de $p < .01$.

Tabla 2. Estadísticos Descriptivos del test “INFVo”

Variables	Media	DT
Inteligibilidad	1,626	,707
Ruido	1,529	,646
Fluidez	1,348	,728
Sonoridad	1,529	,636

También se obtuvieron resultados descriptivos pero esta vez sobre los diferentes parámetros de Vocalab 4, como indica la tabla 3. En ella se aprecia una gran variabilidad entre las puntuaciones registradas por el programa de análisis acústico, al menos en tres de sus parámetros.

Tabla 3. Resultados descriptivos sobre los parámetros de Vocalab 4

Variables	Media	DT
Alteración de ataque	8,166	11,032
Inestabilidad de amplitud	50,766	33,206
Inestabilidad de altura	2,699	1,576
Ruido/señal	1,910	1,445
Pobreza armónica	6,132	17,481

Tras conocer los resultados descriptivos de las variables de estudio, se procedió a realizar una correlación de Pearson entre los parámetros acústicos de Vocalab 4 y los parámetros perceptivos del test “INFVo”. Y por lo que se observa en la tabla 4, se encontraron datos significativos entre las variables perceptivas “Media Inteligibilidad” (,037), “Media Fluidez” (,002), “Media Sonoridad” (,006), “VGP” (,007) y la variable acústica “Ruido/Señal” con una significación $p < ,01$, a excepción de la variable “Media Inteligibilidad” que una obtuvo una $p < ,05$. Por lo que se asume, con un alto nivel de confianza, que existe una fuerte relación entre cuatro de los parámetros perceptivos y el parámetro del programa de análisis acústico, “Ruido/Señal”.

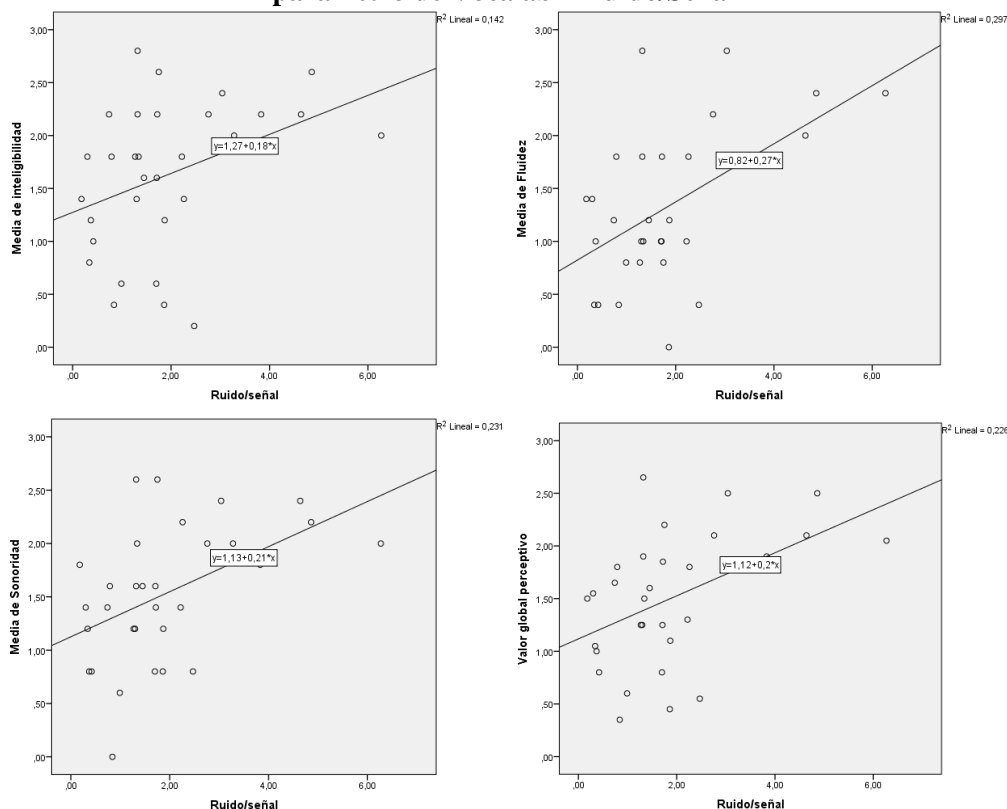
Tabla 4. Coeficientes de correlación de Pearson entre los parámetros de “INFVo” y Vocalab 4

“INFVo”	Vocalab 4				
	Alteración de ataque	Inestabilidad de amplitud	Inestabilidad de altura	Ruido/Señal	Pobreza armónica
Media Inteligibilidad					
Correlación de Pearson	-,127	-,113	-,276	,377*	,176
Sig. (bilateral)	,497	,546	,133	,037	,343
N	31	31	31	31	31
Media Ruido					
Correlación de Pearson	-,117	-,268	-,326	,329	,140
Sig. (bilateral)	,532	,145	,073	,071	,453
N	31	31	31	31	31
Media Fluidez					
Correlación de Pearson	-,036	-,142	-,263	,545**	,245
Sig. (bilateral)	,845	,446	,153	,002	,185
N	31	31	31	31	31
Media Sonoridad					
Correlación de Pearson	-,145	-,205	-,126	,481**	,166
Sig. (bilateral)	,435	,268	,500	,006	,371
N	31	31	31	31	31
VGP					
Correlación de Pearson	-,114	-,196	-,273	,476**	,201
Sig. (bilateral)	,540	,291	,137	,007	,279
N	31	31	31	31	31

* La correlación es significativa en el nivel 0,05. ** La correlación es significativa en el nivel 0,01.

Finalmente, se observó también el tipo de relación que existe entre las variables que presentaron significación, para ello se debe acudir al gráfico 1. En él se observa que existe una relación lineal positiva en todos los diagrama de dispersión. Aunque si se atiende de forma independiente a cada uno de ellos, se observa que la relación que existe entre “Media Inteligibilidad” y “Ruido/Señal” nos indica que a medida que se percibe mejor inteligibilidad, en las voces de nuestra muestra, disminuye el ruido registrado. En cuanto a la relación de la variable “Media Fluidez” se observa lo contrario, aumenta la presencia de ruido cuando disminuye la percepción de fluidez. Coincidiendo con el resto de variables como la “Media Sonoridad” y “VGP”, que cuanto mayor sea el ruido registrado menor será la puntuación recibida tanto en relación a la calidad como a la percepción general de las voces de aperiódicas.

Gráficos 1. Diagrama de dispersión entre las variables significativas del test “INFVo” y el parámetro de Vocalab 4 Ruido/Señal



5. Discusión y conclusiones

El estudio se realizó con el fin de determinar si el programa Vocalab 4 era una buena herramienta para evaluar voces de personas con LT. Para ello, sometimos sus resultados de análisis a los obtenidos por los datos perceptivos, a través del test de evaluación “INFVo”. Con el que pretendíamos conocer si existía alguna relación entre lo registrado por el programa de análisis acústico y el oído humano.

Entre los resultados mostrados hasta ahora debemos analizar en primer lugar los relativos a los parámetros acústicos que no presentaron significación. Dichos resultados muestran una tendencia a presentar valores, en caso de jitter y shimmer, parecidos a los publicados en diferentes estudios. Así, Decore et al. (2001) informa a través de los valores de jitter 2,18% y shimmer 6,77 % que ambos resultados registran grandes alteraciones en voces esofágicas. Al igual que Maccallum, J. K., Cai, L., Zhou, L., Zhang, Y., & Jiang, J. J. (2009), también obtuvieron resultados muy similares de jitter 3,65% y shimmer 47,26%, coincidiendo con los datos obtenidos por Vocalab 4, aunque dicho programa utiliza una

escala de medida diferente para sus puntuaciones, ambos coinciden en la existencia de una gran alteración acústica en estos parámetros analizados. Por tanto, se pone en duda que dichos parámetros sean un buen instrumento para valorar la calidad de las voces aperiódicas debido a su gran variabilidad y poca fiabilidad de los datos obtenidos con el cálculo de jitter y shimmer (Maryn, Roy, De Bodt, Van Cauwenberge, & Corthals, 2009). Respecto al parámetro alteración de ataque se observa que la media de sus puntuaciones es 8,16 con una desviación estándar de 11,32, el cual muestra una gran diferencia entre la puntuaciones de voces aperiódicas con respecto a las voces laríngeas sin patología, las cuales oscilan entre las puntuaciones 0 y 1. Dichos resultados presentan valores esperados, ya que este parámetro se encarga de medir la puesta en vibración de las cuerdas vocales en los primeros 300 milisegundos de su activación. Al conocer que nuestra muestra prescinde de segmentos vocales, no cabría esperar resultados significativos. Conclusión que se obtiene a raíz de conocer la definición del parámetro, ya que no se hallan evidencias sobre la eficacia de su uso para la evaluación objetiva. La incorporación de este parámetro es una medida de carácter innovador en el análisis acústico, ya que ha sido incorporado por los creadores de Vocalab 4 (Menin-sicard & Sicard, 2016). Con respecto al parámetro pobreza armónica, tampoco se obtienen valores significativos. Dichos resultados registran una gran alteración en la energía espectral de los componentes armónicos de la señal recogida (Eliset, 2012). Datos difícilmente comparables con los estudios encontrados, ya que estos no muestran de forma independiente sólo los valores armónicos, sino que son registrados siempre en relación a los componentes de ruido. En este caso Vocalab 4 segmentó el parámetro tradicional HNR en dos parámetros independientes, motivo por el cual no existen evidencias en el análisis acústico que recojan sólo componentes armónicos. Finalmente queda destacar el parámetro ruido/señal siendo este el único que obtuvo significación con cuatro de los cinco parámetros perceptivos ($p < 0,05$). Dichos datos nos informan de que este parámetro es capaz de registrar valores que se relacionan fuertemente con los percibidos por el oído humano.

En el apartado de los resultados relativos a los parámetros perceptivos podemos destacar la variable media ruido, siendo la única variable perceptiva que no obtuvo significación con el parámetro acústico ruido/señal. Dichos datos resultaron anecdóticos, ya que al correlacionar ambas variables, cuyas funciones eran registrar la cantidad de ruido, cabría esperar una fuerte correlación. Por lo que esta falta de efecto significativo se debe a los diferentes criterios de registro del ruido. En cuanto a la variable perceptiva media ruido

cabe destacar la significación obtenida tras la correlación de Pearson con el conjunto de variables perceptivas ($p < .01$), coincidiendo con los datos que obtuvo Moerman (2006) en sus estudios de validez del test.

Finalmente se concluye que el conjunto de medidas registradas por los parámetros de Vocalab 4 no determinan que dicho software sea una buena herramienta de evaluación para voces aperiódicas. Pero sí afirmar que dicho programa, aunque sólo en uno de sus parámetros, sí registra valores de igual modo que el oído humano. Con esto podemos sostener que dicho ajuste de algoritmos, a los que fue sometido esta versión del programa, permite demostrar que se está más cerca de conseguir que los programas de evaluación sobre calidad vocal obtengan mayor precisión de análisis.

6. Referencias bibliográficas

- Cohen, S., & Garrett, C. (2006). Voice-related quality of life in T1 glottic cancer: irradiation versus endoscopic excision. *Annals of Otology*, . Retrieved from <http://aor.sagepub.com/content/115/8/581.short>
- D'Alatri, L., Bussu, F., Scarano, E., Paludetti, G., & Marchese, M. R. (2012). Objective and subjective assessment of tracheoesophageal prosthesis voice outcome. *Journal of Voice : Official Journal of the Voice Foundation*, 26(5), 607–13. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/22209062>
- Deore, N., Datta, S., Dwivedi, R., Palav, R., Shah, R., Sayed, S., ... Kazi, R. (2011). Acoustic analysis of tracheo-oesophageal voice in male total laryngectomy patients. *The Annals of The Royal College of Surgeons of England*, 93(7), 523–527. <http://doi.org/10.1308/147870811X13137608454975>
- Doyle, P. C., & Ph, D. (2015). Auditory-Perceptual Speech Outcomes and Quality of Life after Total Laryngectomy. *Otolaryngol Head and Neck Surg*, 148(July 2012), 82–88. <http://doi.org/10.1177/0194599812461755.Auditory-Perceptual>
- Drugman, T., Rijckaert, M., Janssens, C., & Remacle, M. (2015). Tracheoesophageal speech: A dedicated objective acoustic assessment. *Computer Speech & Language*, 30(1), 16–31. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0885230814000746>
- Eliset, N. (2012). a Coustic Analysis of Normal and Pathological Voices Using Two Different Systems : Anagraf and Praat, 339–358.
- Ferrat, K., & Guerti, M. (2012). A study of sounds produced by Algerian esophageal speakers, 12(4).
- Fomby, P., & Cherlin, A. J. (2011). Coping and Quality of Life After Total Laryngectomy, 72(2), 181–204. <http://doi.org/10.1038/nature13314.A>
- Haderlein, T. (2007). Automatic evaluation of tracheoesophageal substitute voices. Retrieved from <http://www5.informatik.uni->

erlangen.de/Forschung/Publikationen/2007/Haderlein07-AEOb.pdf

- Heuillet-Martin, G. & Conrad, L. (2003). *Hablar sin laringe*. (Lebon, Ed.). Barcelona.
- Jacobi, I., & Timmermans, A. (2015). Voice quality and surgical detail in post-laryngectomy tracheoesophageal speakers. *European Archives of ...* Retrieved from <http://link.springer.com/article/10.1007/s00405-015-3777-4>
- Maccallum, J. K., Cai, L., Zhou, L., Zhang, Y., & Jiang, J. J. (2009). Acoustic analysis of aperiodic voice: perturbation and nonlinear dynamic properties in esophageal phonation. *Journal of Voice : Official Journal of the Voice Foundation*, 23(3), 283–90. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3212416&tool=pmcentrez&rendertype=abstract>
- Mart, R., Barandiar, U., Mart, R., Barandiar, U., & Fern, S. (2006). Voz esofágica, 50, 56–64.
- Maryn, Y., Roy, N., De Bodt, M., Van Cauwenberge, P., & Corthals, P. (2009). Acoustic measurement of overall voice quality: a meta-analysis. *The Journal of the Acoustical Society of America*, 126(5), 2619–2634. JOUR. <http://doi.org/10.1121/1.3224706>
- Menin-sicard, A., & Sicard, E. (2016). Métodos objetivos para la evaluación de la fonación.
- Moerman, M. B. J., Martens, J. P., Van Der Borgt, M. J., Peleman, M., Gillis, M., & Dejonckere, P. H. (2006). Perceptual evaluation of substitution voices: Development and evaluation of the (D)INFVo rating scale. *European Archives of Oto-Rhino-Laryngology*, 263(2), 183–187. <http://doi.org/10.1007/s00405-005-0960-z>
- Ouattassi, N., & Benmansour, N. (2015). Acoustic assessment of erygmophonic speech of Moroccan laryngectomized patients. *The Pan African ...* Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4633833/>
- Sirić, L., Sos, D., Rosso, M., & Stevanović, S. (2012). Objective assessment of tracheoesophageal and esophageal speech using acoustic analysis of voice. *Collegium Antropologicum*, 36 Suppl 2, 111–4. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/23397767>
- Titze, I. R. (1995). Workshop on acoustic voice analysis; Summary Statement. National Center for Voice and Speech, Denver. *Journal of Voice*, 24(3), 29–34. [http://doi.org/10.1016/S0892-1997\(97\)80022-7](http://doi.org/10.1016/S0892-1997(97)80022-7)
- Tóth, A., Csernoch, L., Sziklai, I., & Szűcs, A. (2014). The role of the different neoglottis forms in the development of esophageal voice. *Acta Physiologica Hungarica*, 101(3), 291–300. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/24866929>
- Villarín, M. H. C. (2006). Evaluación acústica y análisis prosódico de la voz esofágica.
- Yan, N., Ng, M. L., Wang, D., Zhang, L., Chan, V., & Ho, R. S. (2013). Nonlinear dynamical analysis of laryngeal, esophageal, and tracheoesophageal speech of Cantonese. *Journal of Voice : Official Journal of the Voice Foundation*, 27(1), 101–10. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/23044459>