

# **ANÁLISIS DISCRIMINANTE APLICADO A LOS FORMANTES DE LAS VOCALES CASTELLANAS EN LA FONACIÓN CON PRÓTESIS Y ERIGMOFONÍA TRAS LARINGUECTOMÍA TOTAL**

**M. ROSIQUE\*, J. L. RAMÓN\*\*, M. CANTERAS\*\*\*, L. ROSIQUE\*\***

\*SERVICIO DE OTORRINOLARINGOLOGÍA. HOSPITAL UNIVERSITARIO VIRGEN ARRIXACA. MURCIA.

\*\* DEPARTAMENTO DE FÍSICA MÉDICA. FACULTAD DE MEDICINA. UNIVERSIDAD DE MURCIA.

\*\*\* DEPARTAMENTO DE BIOESTADÍSTICA. FACULTAD DE MEDICINA. UNIVERSIDAD DE MURCIA.

## **RESUMEN**

**E**l objetivo del presente trabajo consiste en analizar si existen diferencias estadísticamente significativas entre determinadas variables cuantitativas de parámetros acústicos de la voz, en tres muestras de individuos, utilizando análisis discriminante. Los grupos de estudio están constituidos por 20 pacientes con laringuectomía total y voz erigmofoónica, 20 pacientes con laringuectomía total con fistuloplastia y prótesis fonatoria, y 20 individuos normales. Analizamos la energía, frecuencia y ancho de banda de los

cuatro primeros formantes de las cinco vocales castellanas en el contexto de una frase convenida. En función de nuestros resultados el método de análisis discriminante de las variables utilizadas resulta válido para establecer diferencias entre los tres grupos estudiados que, de esta forma, presentan unas características propias desde el punto de vista acústico. Este método nos permite afirmar que la voz traqueoesofágica no se aproxima más a la voz normal que una voz esofágica.

**PALABRAS CLAVE:** Análisis voz. Formantes. Prótesis fonatoria.

## **ABSTRACT**

### **DISCRIMINANT ANALYSIS APPLIED TO THE FORMANTS OF THE CASTILIAN VOWELS IN THE PHONATION WITH PROSTHESIS AND ESOPHAGEAL VOICE AFTER TOTAL LARINGECTOMY**

**T**he objective of the present work consists, in analyzing if there is statistically significant differences between certain quantitative variables of acoustic parameters of the voice, in three samples of individuals, using a discriminative analysis. The study groups are formed by: 20 patients with total laryngectomy and esophageal voice, 20 patients with total laryngectomy and tracheoesophageal voice with phonatory prosthesis, and 20 normal individuals. We analyzed the energy, bandwidth and fre-

quency of the four first formants of the five Castilian vowels in the context of an established phrase. Based on our results the method of discriminative analysis of the used variables is valid to establish differences between the three studied groups; that is they display their own characteristics from the acoustic point of view. This method allows us to affirm that the tracheoesophageal voice does not come closer to the normal voice that the esophageal voice does.

**KEY WORDS:** Voice analysis. Formants. Phonatory prosthesis.

Este trabajo ha sido financiado por el Fondo de Investigaciones Sanitarias de la Seguridad Social (FISS), Exped. 97/1153.

**Correspondencia:** Mariano Rosique Arias. Carril Salabosque, 3. 30012 Murcia.

**Fecha de recepción:** 10-2002

**Fecha de aceptación:** 25-2-2003

## INTRODUCCIÓN

El sonido generado en la laringe que origina la voz por sí solo no sirve para la comunicación oral, por lo que debe sufrir una serie de transformaciones en las cavidades supralaríngeas y llegue a ser un sonido inteligible el que salga por los labios. Si consideramos la laringe como fuente sonora, las estructuras referidas por encima de ella se denominan tracto vocal. El tracto vocal está constituido por la laringe supraglótica, la faringe, la boca y las fosas nasales. El tracto vocal, durante el habla, está modificando continuamente su anatomía en la forma de distintos gestos articulatorios de tal manera que se intensifican o atenúan diferentes sonidos. Las zonas donde se intensifican determinados sonidos se llaman resonancias, por lo que cada forma que adquiere el tracto vocal tiene una determinada curva de resonancia. La mecánica del habla necesita el cambio continuo de curvas de resonancia y de las fuentes sonoras acopladas a ella para producir los diferentes fonemas del habla<sup>1-3</sup>. Si deseamos emitir una vocal determinada imponemos al tracto vocal una cierta conformación que lo convierte en un resonador complejo a través del cual se propaga el sonido laríngeo<sup>4-6</sup>. Así pues, en los espectros sonoros de las vocales existen grupos de armónicos de diferentes rangos de frecuencia, que presentan máximos relativos de intensidad, y corresponden a los formantes, distinguiéndose hasta un quinto formante de acuerdo con los valores crecientes de sus frecuencias medias. Estas frecuencias medias para cada vocal no son valores rígidos, pueden variar de locutor en locutor y aún dentro de cada uno de ellos, de todas formas el rango de variación de las frecuencias donde aparecen los formantes tiene unos límites, sobrepasados los cuales, la inteligibilidad del habla corre el riesgo de deteriorarse. Según Sundberg<sup>7</sup> hay formas de modificar el tracto vocal y por lo tanto la frecuencia de los formantes; por ejemplo, si se desciende la laringe o se produce un entubamiento de los labios se puede alargar el tracto vocal consiguiéndose un descenso del primer formante, que dará como resultado una voz más oscura. La maniobra opuesta produce una elevación del primer formante.

Las características de los formantes son la base en la comparación e identificación de diferentes voces<sup>8,9</sup>, y especialmente las frecuencias altas del espectro aportan al receptor información acústica sobre el timbre personal del sonido de la voz<sup>10</sup>. Por ejemplo el tercer y cuarto formantes son puntos muy significativos en la identificación de las personas ya que los formantes más altos son producto

del timbre individual de cada locutor<sup>11</sup>. Estos picos o frecuencias formánticas son debidos no necesariamente a un armónico, sino a un grupo de armónicos de frecuencias próximas a la de resonancia<sup>12</sup>. Un formante se define como un máximo en la función de transferencia del tracto vocal<sup>13,14</sup>. Cada vocal se caracteriza por unos formantes determinados que distinguen unas de otras, así las vocales castellanas se pueden clasificar en función de su estructura armónica o timbre en vocales agudas (e, i), vocales graves (o, u), y neutras (a). Mediante análisis LPC (*Lineal Predictive Coding*) del que disponen la mayoría de los programas informáticos analizadores de voz, podemos identificar los valores en frecuencia de los distintos formantes, su energía o intensidad y su definición o grado de aplanamiento mediante el valor ancho de banda (*bandwidth*).

El estudio acústico aplicado a las voces patológicas se basa sobre todo en la cuantificación de la cantidad de ruido indeseable en la señal acústica, o en la alteración de la periodicidad en la vibración de las cuerdas vocales<sup>15,16</sup>. Para ello se utilizan programas informáticos que mediante algoritmos matemáticos proporcionan parámetros que definen esas alteraciones. Esto supone una ayuda en el diagnóstico e incluso una valoración más de los resultados fonatorios tras una microcirugía endolaríngea, o la cirugía parcial de laringe<sup>17-19</sup>.

## HIPÓTESIS

Si podemos considerar los formantes como característicos del tracto vocal de cada persona, podemos suponer que también definen un tracto modificado tras la laringectomía total. Mediante los parámetros frecuencia, energía y ancho de banda que definen los distintos formantes, podemos estructurar y diferenciar una voz normal de una voz esofágica y traqueoesofágica. Este es el punto de partida que utilizaremos como hipótesis de trabajo.

## MATERIAL Y MÉTODO

Desde 1990 realizamos en nuestro Servicio de forma protocolizada cirugía fonatoria tras laringectomía total, mediante la confección de una fístula traqueoesofágica y colocación posterior de prótesis fonatoria. Para nuestro estudio hemos seleccionado 20 pacientes portadores de prótesis fonatoria de Herrmann de silicona y acero inoxidable con curvatura de 90°, que se colo-

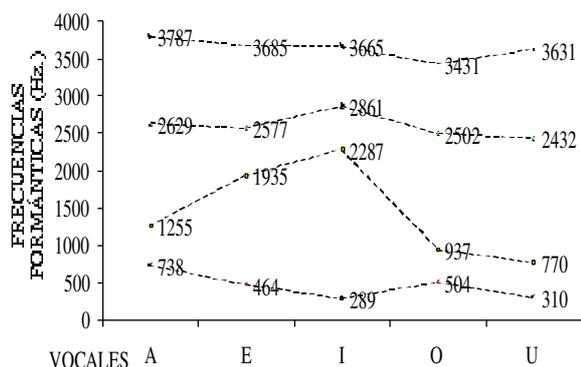


Figura 1. Valores de los formantes con voz normal.

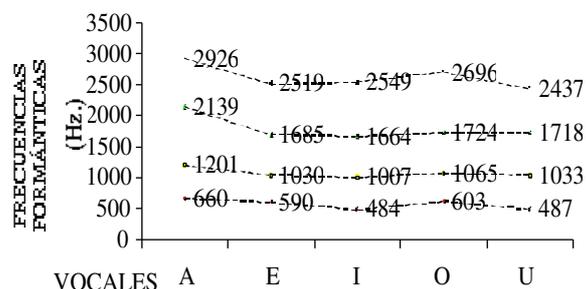


Figura 2. Valores de los formantes con prótesis fonatoria.

có tras laringuectomía total con fistuloplastia primaria y cierre de faringe con colgajo dermo-platismo-fascial<sup>20</sup>. A todos estos pacientes se les realizó miotomía cricofaríngea. Otros 20 enfermos con voz erigimofónica habían sufrido laringuectomía total y cierre faríngeo simple con aproximación de bordes con el mismo tipo de sutura. Todas las neoformaciones correspondían a estadio T3 glótico. Estos enfermos habían sido calificados con buena voz ante un auditorio de jueces normooyentes que realizaron un análisis perceptual de la voz<sup>21</sup>. Los dos grupos no mostraban diferencias estadísticamente significativas en cuanto a variables de tipo clínico ni edad. Desde el punto de vista oncológico ningún sujeto presentaba signos de recidiva tumoral. Como grupo control utilizamos 20 individuos de sexo masculino con audición y fonación normal, comparables en la variable edad a los anteriores.

Todos los sujetos del estudio pronunciaron la misma frase tres veces para su grabación en una cámara insonorizada. El sistema de grabación y las condiciones fueron similares en todos los casos. La frase utilizada fue: "la lluvia en Sevilla no es una maravilla". Las vocales para estudio fueron la /a/ de la, la /e/ de es, la /i/ de maravilla, la /o/ de no, y la /u/ de lluvia. El análisis acústico se hizo con el programa Multi speech 3700 de Kay Elemetrics, identificando los formantes de las vocales descritas con LPC. Hemos obtenido así los cuatro formantes de las cinco vocales con sus correspondientes intensidades y anchos de banda lo que hace un total de 60 valores numéricos para cada muestra de voz obtenida de los diferentes sujetos.

Para el análisis estadístico se empleó el paquete estadístico BMDP Statistical Software Inc. 1988 (IBM PC/DOS). El estudio multivariante se realizó con análisis discriminante. Mediante este

procedimiento se pueden determinar qué variables son las que más diferencian a la hora de clasificar a un sujeto en distintos grupos. Se obtienen funciones discriminantes que son combinaciones lineales de dichas variables. Para un caso determinado la función discriminante tomará un valor concreto, y dicho caso se clasificará en uno u otro grupo, atendiendo a cuál sea el mayor valor resultante de dicha función.

## RESULTADOS

En las figuras 1, 2 y 3 se muestran los valores medios observados de las frecuencias de los cuatro formantes vocálicos. Corresponden a los tres grupos de individuos normales (Figura 1), con prótesis fonatoria (Figura 2) y erigimofonía (Figura 3).

A continuación mostramos las amplitudes medias en decibelios de los distintos formantes de las cinco vocales en las tres muestras de individuos (Tablas 1, 2 y 3).

La anchura de banda de cada formante se expone en las Tablas 4, 5 y 6.

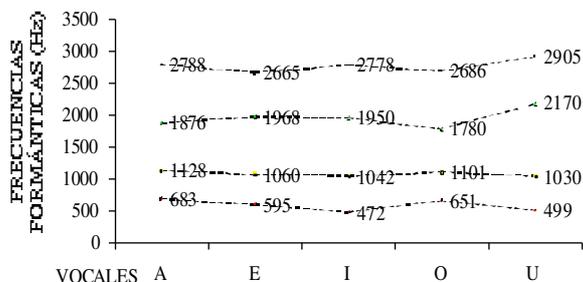


Figura 3. Valores de los formantes con erigimofonía.

**Tabla 1: Amplitudes en decibelios de los formantes en el grupo con voz normal**

	A	E	I	O	U
A1	46	57	65	52	60
A2	54	61	64	50	57
A3	52	55	53	56	57
A4	52	51	48	57	56

**Tabla 2: Amplitudes en decibelios de los formantes en el grupo con prótesis fonatoria**

	A	E	I	O	U
A1	47	52	65	52	57
A2	55	57	63	50	52
A3	57	55	63	52	64
A4	54	52	49	53	64

**Tabla 3: Amplitudes en decibelios de los formantes en el grupo con erigmo fonía**

	A	E	I	O	U
A1	45	48	54	48	54
A2	51	54	57	50	53
A3	48	54	57	53	54
A4	42	51	47	55	54

**Tabla 4: Ancho de banda en hertzios en el grupo con voz normal**

	A	E	I	O	U
B1	32	51	96	70	75
B2	118	133	125	90	111
B3	102	123	109	154	150
B4	157	114	104	160	120

La aplicación del análisis discriminante a todas las variables obtenidas permite obtener los valores del estadístico F para las diferentes vocales siendo todos significativos. En las siguientes tablas mostramos estos resultados (Tablas 7, 8, 9, 10 y 11).

De forma esquemática representamos en la Figura 4 las distintas áreas donde se ubican los tres grupos estudiados. A pesar de que existe cierto solapamiento cada individuo tiene tendencia a situarse en su área correspondiente.

## DISCUSIÓN

Los resultados obtenidos en el análisis discriminante permiten diferenciar los tres grupos

por las características de las variables estudiadas, aunque no podemos afirmar que la voz en los portadores de prótesis fonatoria ni la erigmo fonía se aproximan a la voz normal en mayor medida una u otra. Si en el análisis discriminante hubiésemos encontrado la inexistencia de diferencias estadísticamente significativas entre, por ejemplo, la voz traqueo-esofágica y la voz normal, se habría podido concluir que la voz con prótesis fonatoria se aproxima más a la normal con los parámetros que hemos estudiado, hecho que no ha ocurrido así, en contra de lo admitido en la bibliografía<sup>21,22</sup> cuando se hace un análisis perceptual de las voces ante un auditorio. Es posible que con el avance futuro de programas de análisis de voz capaces de determinar nuevos

**Tabla 5: Ancho de banda en hertzios en el grupo con prótesis fonatoria**

	A	E	I	O	U
B1	40	55	148	52	70
B2	100	97	144	101	70
B3	137	88	207	164	205
B4	134	133	93	133	169

**Tabla 6: Ancho de banda en hertzios en el grupo con erigmo fonía**

	A	E	I	O	U
B1	44	55	85	65	89
B2	117	109	113	61	81
B3	140	191	172	123	185
B4	203	164	103	124	155

**Tabla 7: Valores de F en análisis discriminante al comparar los tres grupos de pacientes para la vocal "a". Todas las cifras son significativas**

vocal "a"	Prótesis fonatoria	Erigmofofía
Erigmofofía	8,41	
Normales	8,18	5,94

**Tabla 8: Valores de F en análisis discriminante al comparar los tres grupos de pacientes para la vocal "e". Todas las cifras son significativas**

vocal "e"	Prótesis fonatoria	Erigmofofía
Erigmofofía	6,76	
Normales	12,28	7,70

**Tabla 9: Valores de F en análisis discriminante al comparar los tres grupos de pacientes para la vocal "i". Todas las cifras son significativas**

vocal "i"	Prótesis fonatoria	Erigmofofía
Erigmofofía	8,43	
Normales	10,53	6,34

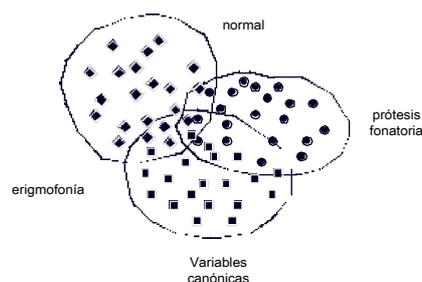
**Tabla 10: Valores de F en análisis discriminante al comparar los tres grupos de pacientes para la vocal "o". Todas las cifras son significativas**

vocal "o"	Prótesis fonatoria	Erigmofofía
Erigmofofía	5,43	
Normales	8,83	1,79

parámetros se puedan establecer diferencias cuantitativas más específicas de las obtenidas con los medios actuales. La vibración tanto en la erigmofofía como en la voz traqueoesofágica se produce en el mismo lugar, los repliegues mucosos que constituyen la pseudoglótis fonatoria del esófago en la ubicación del músculo cricofaríngeo. La discriminación posiblemente se establezca entre las muestras de laringectomizados por las diferentes técnicas quirúrgicas y la morfología del segmento faringoesofágico en cada grupo, más amplio y resonante el colgajo de Herrmann. Obtenemos tres grupos diferenciados como si las características de la voz de cada uno de ellos fueran propias de la técnica quirúrgica practicada y de una voz normal. Este método de estudio es apropiado porque clasifica a cada individuo en su grupo correspondiente. Es decir, con una determinada muestra de voz y analizando las variables que hemos utilizado, podemos ser capaces de determinar en qué patología se ubica.

**Tabla 11: Valores de F en análisis discriminante al comparar los tres grupos de pacientes para la vocal "u". Todas las cifras son significativas**

vocal "u"	Prótesis fonatoria	Erigmofofía
Erigmofofía	14,48	
Normales	6,07	14,24



**Figura 4. Esquema simplificado de los tres grupos estudiados en el análisis discriminante de la voz.**

## REFERENCIAS

- White P. Formant frequency analysis of children's spoken and sung vowels using sweeping fundamental frequency production. *J Voice* 1999; 13(4): 570-82.
- Cheveigne A, Kawahara H. Missing data model of vowel identification. *J Acoust Soc Am* 1999; 105(6): 3497-508.
- Hosemann W, Gode U, Dunker J E, Eysholdt U. Influence of endoscopic sinus surgery on voice quality. *Eur Arch Otorhinolaryngol* 1998; 255(10): 499-503.
- Riede T, Fitch T. Vocal tract length and acoustics of vocalization in the domestic dog. *J Exp Biol* 1999; 202: 2859-67.
- Story B H, Titze I R, Hoffman E.A. Vocal tract area functions for an adult female speaker based on volumetric imaging. *J Acoust Soc Am* 1998; 104(1): 471-87.
- Fitch W T. Vocal tract length and

- formant frequency dispersion correlate with body size in rhesus macaques. *J Acoust Soc Am* 1997; 102: 1213-22.
- 7.- Sundberg J. The acoustics of the singing voice. *Sci Amer* 1977; 82-91.
- 8.- Ramón J, Sprekelsen C, Marin M, González M, Rubio M. Contribución al estudio de la fonación mediante análisis secuencial de frecuencias. *Acta Otorrinolaringol Esp* 1980; 31: 577-88.
- 9.- Tosi O. Elementos de acústica y sicoacústica. En: J Menaldi (ed.). *La voz normal*. Panamericana. Buenos Aires. 1993; 27-45.
- 10.- Martí J. Parámetros vocálicos del catalán. *Folia Fonética* 1984; 1: 22-43.
- 11.- Quilis A. *Curso de fonética y fonología españolas*. Madrid 1982. Consejo Superior de Investigaciones Científicas.
- 12.- Dunn H.K. The calculation of vowel resonances and an electrical vocal tract. *J Acoust Soc Am* 1950; 22: 740-753.
- 13.- Fant G. The vocal tract. En: Fant G. (ed.). *Acoustic theory of speech production*. Mouton. The Hage. Paris. 1970; 13-20.
- 14.- Lienard J. Les procesus de la communication parlée. *Introduction a l'analyse et synthese de la parole*. Paris. Masson. 1977.
- 15.- Dejonckere P, Lebacqz J. Acoustic, perceptual, aerodynamic and anatomical correlations in voice pathology. *J Voice* 1996; 58(6) 326-32.
- 16.- Rosique M, Ramón J, Campos M, Cervantes J. Acoustic analysis of the voice in phonatory fistuloplasty after total laryngectomy. *Otolaryngol Head and Neck Surg* 2000; 112(5): 743-7.
- 17.- Szmeja Z, Leszczynska M. Voice function in patients after extended fronto-lateral laryngectomy. *Eur Arch Otorhinolaryngol* 1999; 256(8): 418-22.
- 18.- Kosztyla-Hojna B, Chodyncki S, Lazarczyk B. The voice function after horizontal laryngectomy. *Otolaryngol Pol* 1998; 52(2): 195-8.
- 19.- Robbins J, Fisher H B, Logemann J A. Acoustic characteristics of voice production after Staffieri's surgical reconstructive procedure. *J Speech Hear Disord* 1982; 47(1): 77-84
- 20.- Herrmann I, Buchwald J, Zenner H. Die glottoplastik-eine neue methode zur chirurgischen stimmrehabilitation. *HNO* 1983; 32: 124-9.
- 21.- Vea J, Perelló J, Algaba J, Zulue-ta A. Estudio de la voz en fistuloplas-tias fonatorias. *An ORL Iber Amer* 1985; 12: 569-75.
- 22.- Max L, Steurs W, Bruyn W. Vocal capacities in esophageal and tracheo-esophageal speakers. *Laryngoscope* 1996; 106: 93-96.