

# Caracterización articulatoria de los sonidos vocálicos del español de Buenos Aires mediante técnicas de resonancia magnética.

Gurlekian, Jorge A., Elisei, Natalia y Eleta, Martín

## I. INTRODUCCION

La caracterización articulatoria de los distintos sonidos del Español de la Argentina ha sido una tarea pendiente desde los primeros registros radiográficos realizados para esta variante del Español. Conocer la forma y dimensiones que éste adopta durante la producción es esencial para comprender los procesos articulatorios y acústicos involucrados. Distintos trabajos han mencionado la necesidad de contar con descripciones articulatorias más precisas y las técnicas de resonancia magnética permiten en la actualidad una observación de los órganos blandos sin producir riesgos debidos a la exposición a la radiación.

Los estudios tradicionales han sido realizados mediante técnicas radiográficas (Criba y Kajiyama, 1941; Fant, 1960; Heinz y Stevens, 1964; Perkell, 1969, Sundberg, 1969). En ellos se describen las desventajas resultantes de la exposición a radiación y la dificultad para deducir con precisión las funciones de área aplicando transformaciones a las medidas obtenidas en el plano sagital.

Otros métodos de diagnóstico que han provisto buena resolución temporal son: Los rayos X *Microbeam* (Fujimura, 1991; Perrier y otros, 1992; Beautemps y otros, 1995; Badin y otros, 1997;), el *ultrasonido* (Stone, 1990), la *electropalatografía* (Nguyen y otros, 1996; Mair y otros, 1996) y la *articulometría electromagnética*

(Schönle, 1987; Perkell, 1992; Hoole y otros, 1993). Las limitaciones en este caso es que las técnicas utilizadas miden el tracto vocal en 2D o en regiones restringidas.

Actualmente existen métodos que brindan información completa en 3D como la *tomografía computada (TC)* basada en rayos X y las *imágenes por resonancia magnética (IRM)*. La TC no es usada en estudios de producción del habla dado que el registro de un sólo fonema sobrepasa el límite de radiación permitido con fines no médicos. La IRM es considerada una técnica confiable y apropiada para medir e investigar en forma estática e inocua la morfología del tracto vocal en 3D durante la fonación a través de imágenes de alta calidad que detallan los tejidos blandos involucrados en la producción del habla como labios, velo del paladar, glotis, etc. Las desventajas de la técnica de adquisición de imágenes por resonancia magnética son: 1) requiere tiempos de barrido mayores –del orden de minutos-, 2) el ruido ambiente durante el registro de las imágenes por IRM y la posición supina del sujeto en el resonador son inevitables y 3) carece de definición para tejidos con bajo contenido de hidrógeno móvil como el óseo y el dentario.

Debido al costo y al tiempo que demandan las IRM, hasta ahora se han realizado en un número reducido de sujetos y lenguas centradas en el estudio de las vocales (Baer y otros, 1991; Yang

y Kasuya, 1994; Demolin y otros, 1996; Story y otros, 1996) y grupos de consonantes (nasales: Matsumura y otros, 1994; Dang y otros, 1994; laterales: Narayanan y otros, 1997 y fricativas: Engwall y Badin, 1999).

Para la variante del español de la Argentina no existen estudios con IRM. Disponemos solo de estudios radiográficos de las vocales españolas (Quilis, A, 1993; Toledo y García Jurado, 1981) y palatogramas (Navarro Tomás, 1918).

Nuestro objetivo principal es describir las vocales del español de la Argentina mediante la técnica de IRM. Luego, teniendo en cuenta que la teoría acústica de la producción de habla considera al tracto vocal como un conjunto de tubos de distinta longitud y con variaciones de la sección de área a lo largo del mismo, se medirán las longitudes en los cortes sagitales, y las áreas del tracto vocal en los cortes coronales, oblicuos y axiales para las cinco vocales emitidas en forma aislada y sostenida por un hablante del español de la Argentina. Posteriormente a partir de estas mediciones se estimaran los valores de frecuencia de los formantes y se los comparará con los producidos acústicamente por el mismo sujeto.

El trabajo esta organizado de la siguiente forma: El equipamiento y los parámetros utilizados para la medición por IRM se describen en la Sección II. Las imágenes obtenidas y su medición se presentan en la Sección III.

## II DESARROLLO EXPERIMENTAL

Las IRM se tomaron con un equipo Philips Gyroscan Intera MR 1.5 con una bobina Tesla Brain en un sujeto femenino entrenado de 26 años de edad, hablante

del español de la Argentina sin antecedentes de patología de voz, habla ni audición. Se verificó que el sujeto tuviese simetría odonto-mandibular.

Las tomas consisten en 3 series de 15 cortes (6 cortes axiales, 6 coronales, 3 coronales-oblicuos) siguiendo los parámetros consignados en la tabla.

**Tabla I. Parámetros seleccionados para la medición por IRM**

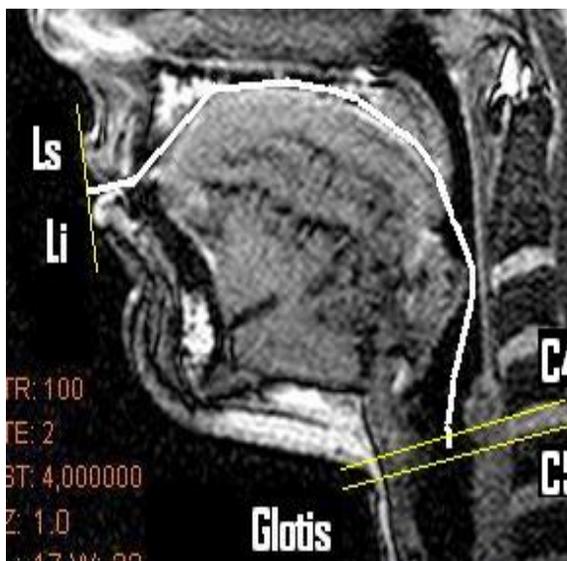
<b>Tiempo de repetición</b>	TR:20 ms	
<b>Tiempo de retardo</b>	TE: 2.0 ms	
<b>Matriz de la imagen</b>	256x256	
<b>Resolución</b>	512	
<b>Campo visual</b>	250 x 200	230 x 230
<b>Espesor del corte</b>	4 mm	4 mm
<b>Factor distancia</b>	5 mm	0 mm
<b>Número de cortes</b>	178	

Se ubicó al sujeto dentro del equipo en posición supina y se alineó la cabeza de manera que el plano de Frankfurt fuese perpendicular a la tabla de imagen y quedase anatómicamente centrada en la línea media. Las órdenes para la secuencia de producción se realizaron a través de un intercomunicador. El sujeto fue entrenado para pronunciar y sostener cada una de las cinco vocales /a, i, e, o, u/ por un lapso de 45 seg. Las emisiones se produjeron con pausas de 30 seg.

Se registraron acústicamente las vocales en tres condiciones: las originales hechas por el sujeto desde el resonador y las realizadas posteriormente en posición supina y sedente.

Una vez tomadas las RM, las imágenes se guardaron en el formato original (DICOM). Para la medición de las áreas de las vocales se utilizó el programa de Philips anexo al equipo de RM. Sin embargo fue necesario transformar las imágenes del formato DICOM a TIFF para medir las longitudes. Se utilizó el programa NHI-Image desarrollado por National Institutes of Health de USA. Este software que está disponible de

manera gratuita en Internet y cuenta con la ventaja de una herramienta “free line” que permite medir la longitud el tracto vocal sin las limitaciones del primero. Se fijaron criterios de medición respecto al trazado de límites (glotis-labios) y estructuras a considerar. La longitud del tracto vocal se definió como la distancia a lo largo de la línea media del tracto vocal, partiendo de una línea recta imaginaria



**Figura 1.** Determinación de la longitud del tracto vocal en un corte sagital de referencia.

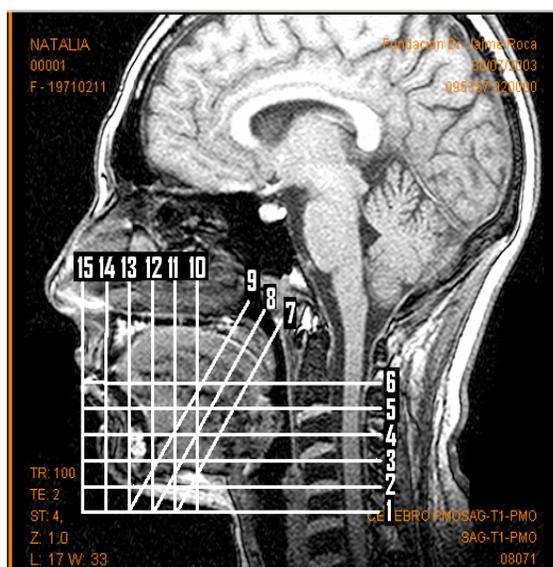
que une las prominencias externas de cada labio (LS y LI) hasta el punto medio de intersección de una recta perpendicular que cruza la tráquea al nivel de un punto medio entre C4 y C5. Figura 1.

Para la medición de las áreas de los primeros 6 cortes transversales o axiales asociados al piso de la laringe, el vestíbulo laríngeo y porción inferior de la faringe no se consideraron los senos piriformes (Baer, 1991) y se descontó del área total el valor de área correspondiente a la epiglotis.

En los 3 cortes coronales oblicuos que investigan la región orofaríngea, la imagen del velo se vio afectada en su calidad a causa de su movimiento durante

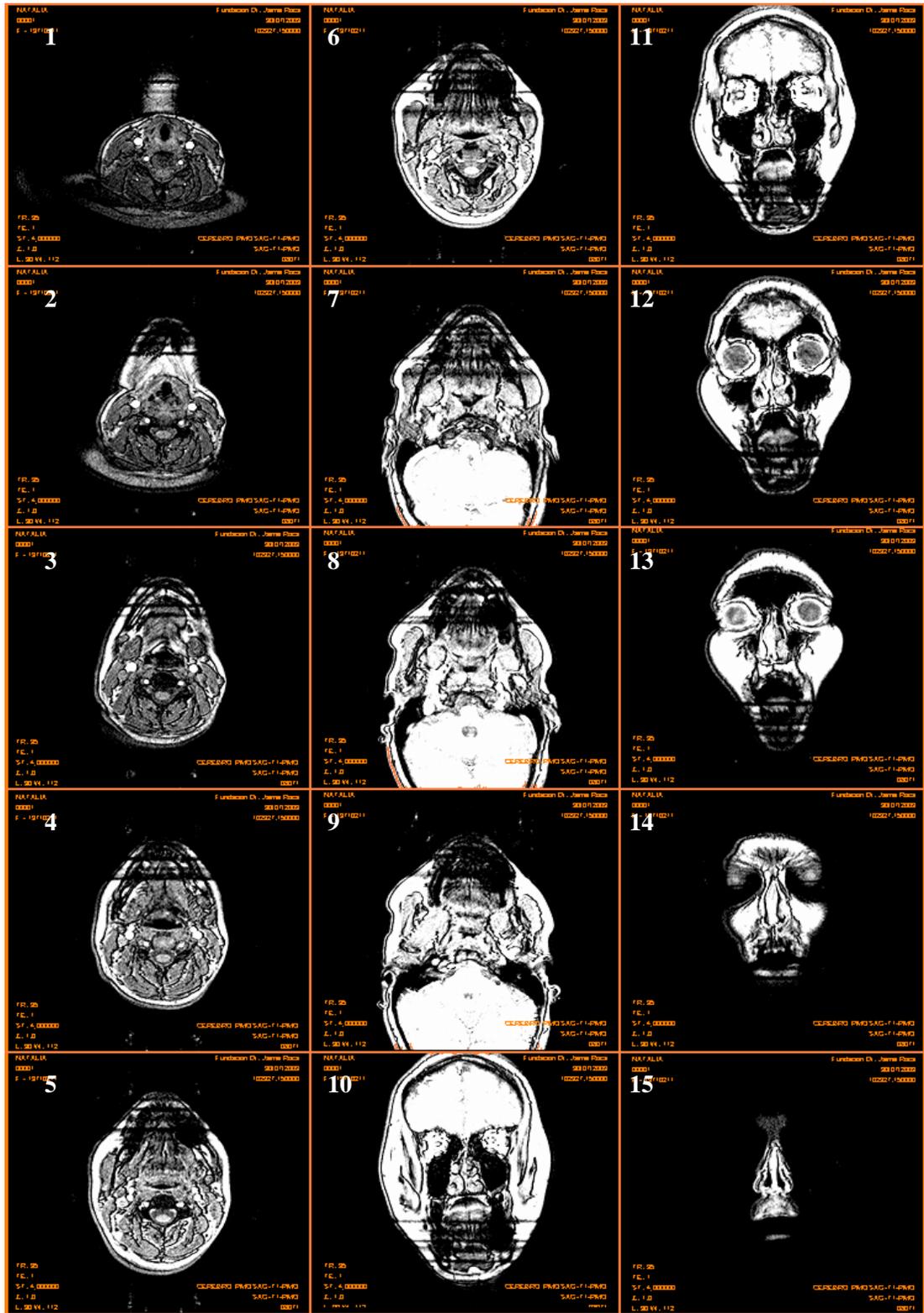
la fonación, la respiración y el flujo sanguíneo (arteria carótida).

En los restantes 6 coronales que muestran las regiones correspondientes a labios, alvéolos y paladar duro, el mayor inconveniente lo produjo la falta de visualización de las piezas dentarias por ser tejidos poco sensibles a los estudios de resonancia magnética.

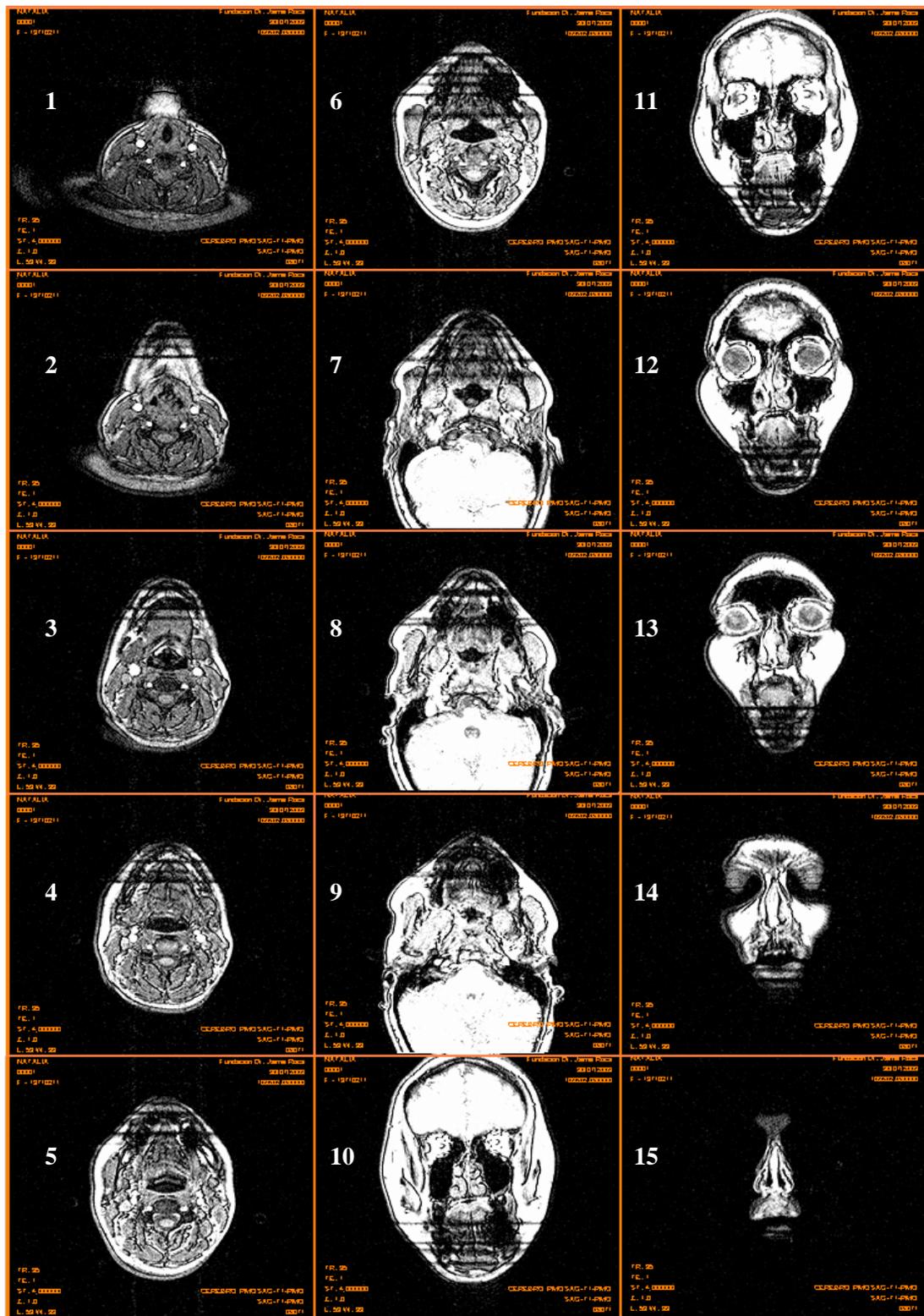


**Figura 2.** Distribución de los cortes coronales (1 a 6) coronales oblicuos, (7 a 9) y transversales (10 a15).

A continuación las Figuras 3,4,5,6 y 7 muestran los cortes para las vocales /i, e, a, o/ y /u/ respectivamente. Luego, la Figura 8 muestra los cortes sagitales para las cinco vocales. En la Figura 9 se presentan las mediciones de longitudes y secciones de área para las vocales /i, a, u/ obtenidas de las figuras anteriores.

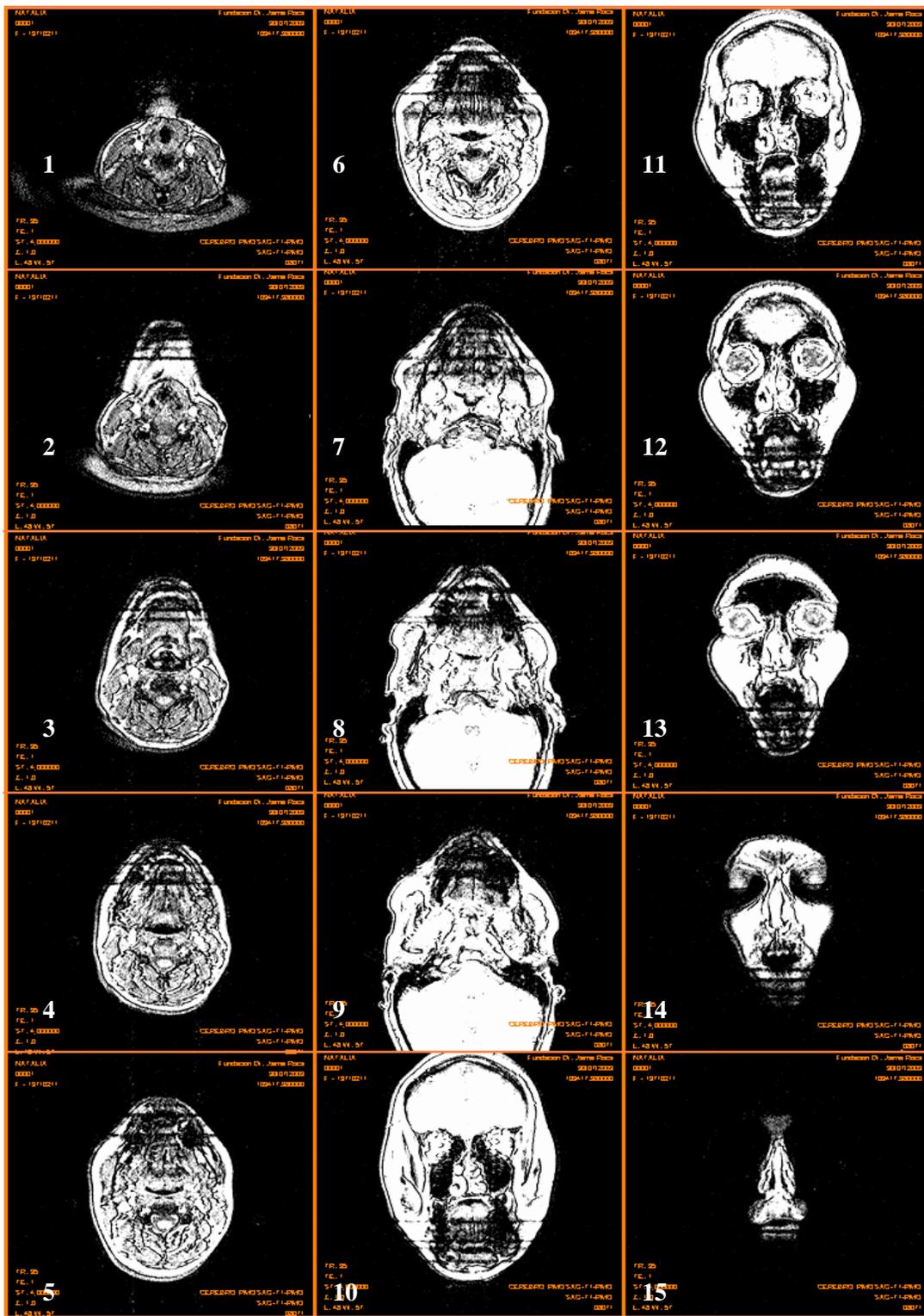


**Figura 3.** Cortes Coroneles (1 a 6), Coroneles oblicuos, (7 a 9) y Transversales (10 a 15) para la vocal a.

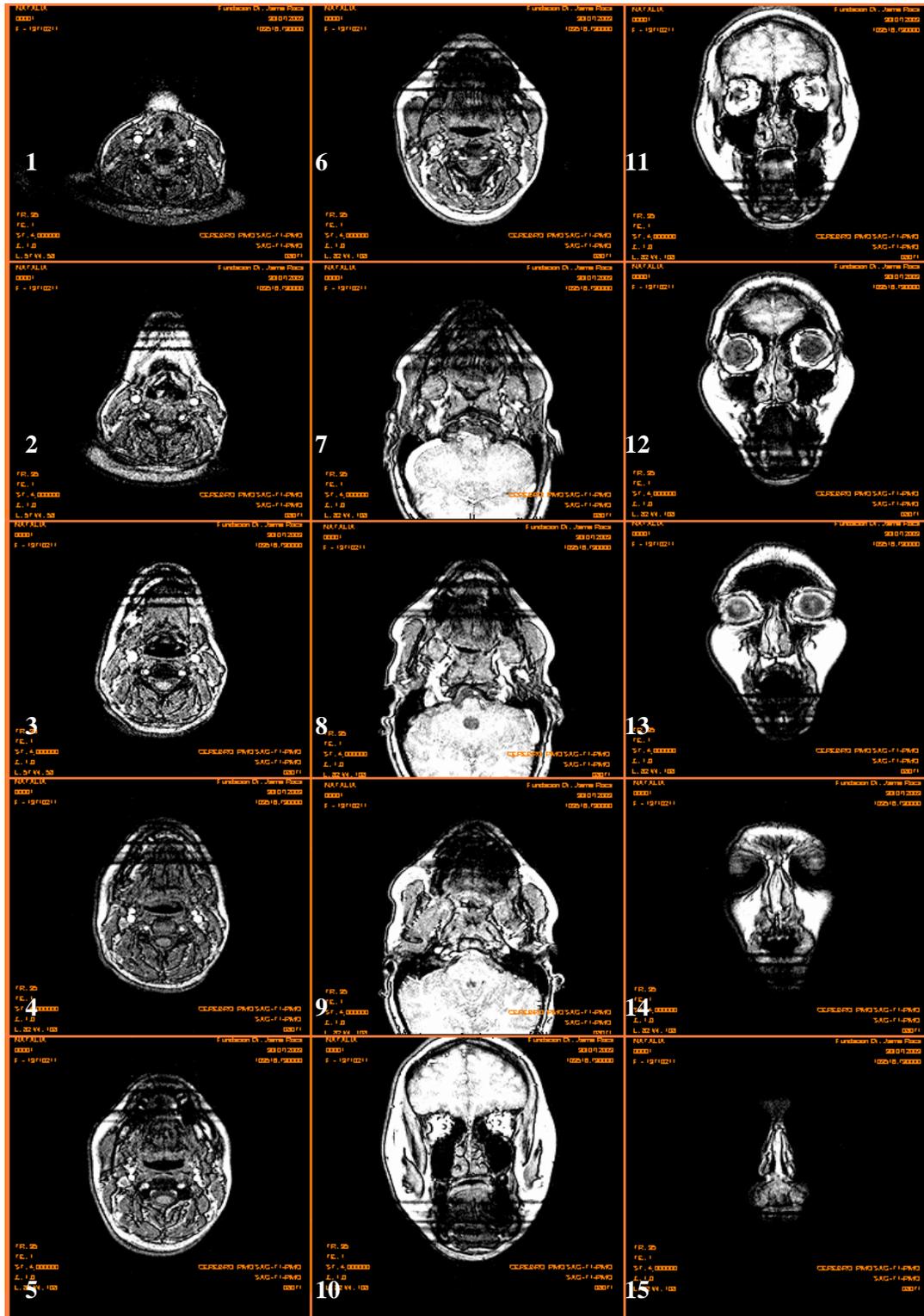


**Figura 4.** Cortes Coroneles (1 a 6), Coroneles oblicuos, (7 a 9) y Transversales (10 a 15) para la vocal i.

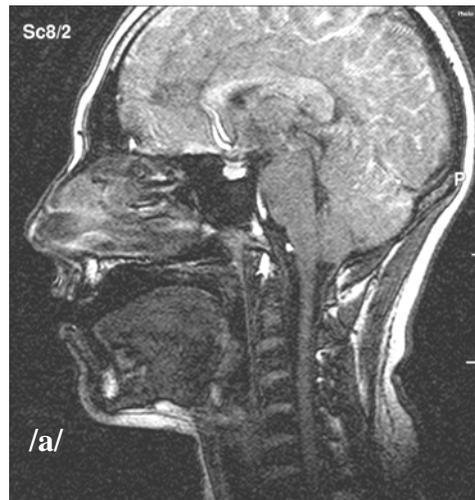
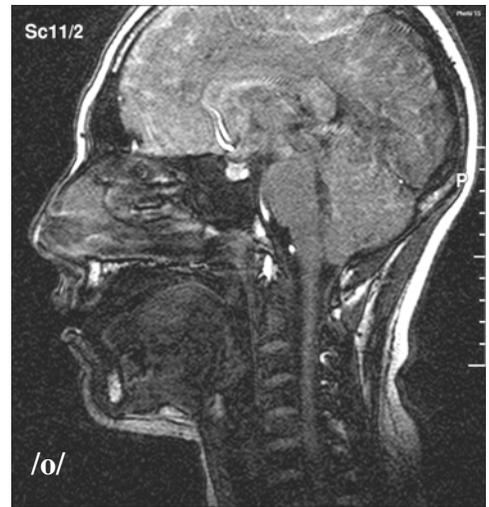
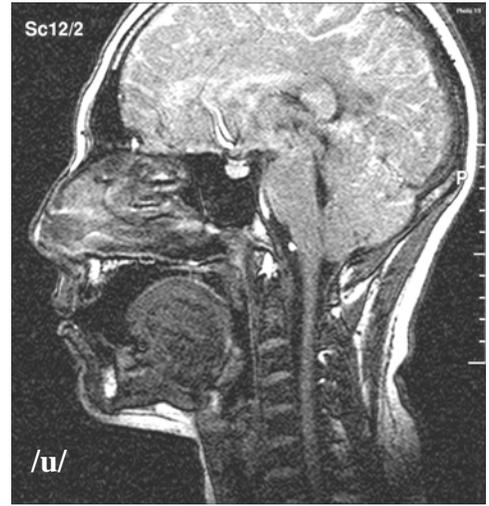
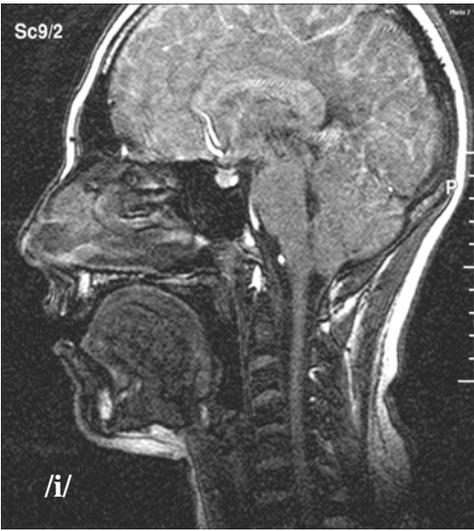




**Figura 6.** Cortes Coronales (1 a 6), Coronales oblicuos, (7 a 9) y Transversales (10 a 15) para la vocal o.



**Figura 7.** Cortes Coronales (1 a 6), Coronales oblicuos, (7 a 9) y Transversales (10 a 15) para la vocal u.

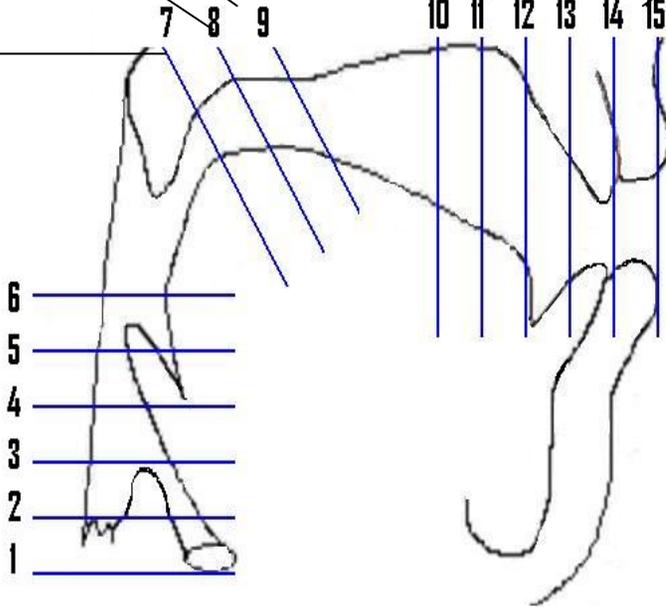


**Figura 8.** Cortes Sagitales para las cinco vocales.

**Figura 9.** Sección de área observada y su medición (A) en cm<sup>2</sup>, correspondiente a los cortes indicados en el esquema sagital del tracto vocal para las longitudes (L) indicadas en cm desde la glotis.

	L	/u/	A	/i/	A	/a/	A
<b>9</b>	8,7		0.34		1.89		0.71
<b>8</b>	7,9		0.71		2.45		0.72
<b>7</b>	7.0		0.58		4.78		2.36
<b>6</b>	5.4		2.55		5.12		1.58
<b>5</b>	4.5		2.95		4.98		1.23
<b>4</b>	3.6		2.61		5.55		2.53
<b>3</b>	2.7		6.23		5.03		5.52
<b>2</b>	1.8		(5.63) 3.02		(4.24) 2.2		(3.27) 1.92

	10	11	12	13	14	15
L	10,3	11,2	12,1	13	13,9	14,8
/u/						
A	0.75	2.19	6.47	6.28	3.09	0.53
/i/						
A	0.30	0.36	0.55	0.43	2.06	1.68
/a/						
A	1.75	2.83	3.99	5.02	4.52	4.88



Las zonas oscuras representan espacios de aire o tejidos con muy baja densidad de moléculas de hidrógeno como los dientes o huesos. Los tejidos grasos y medula ósea aparecen como las más claras, los músculos y tejidos conectivos se observan en distintos niveles de gris. Los vasos sanguíneos aparecen en colores oscuros pues la sangre, aunque rica en hidrógeno, circula velozmente. Como se mencionó esta condición produce cierta indefinición en la visualización de los contornos al igual que la respiración.

El corte #1, realizado por debajo de la glotis muestra la sección de traquea adyacente que no se considerará para el cálculo de las resonancias del tracto vocal. En los cortes #2 y #3 se observan las secciones de área de los senos piriformes a ambos lados y hacia atrás del tubo laríngeo. En el corte #2 los senos están más separados y en la Figura 8 se presenta entre paréntesis la contribución de los senos a la sección de área. Fant (1960) ha indicado que el efecto acústico principal de estas dos cavidades es producir anti-resonancias por encima de los 5kHz. Por este motivo no se considerará su contribución para el cálculo de las frecuencias de formantes, aun cuando su contribución puede reducir levemente la frecuencia de formantes en vocales posteriores.

La epiglotis se observa en la parte superior de la sección de área de los cortes coronales 3, 4 y 5 en las vocales /a, i, o/ y en el corte 4 en la vocal /u/. En los sonidos que producen una constricción en esta zona /a/ la epiglotis se adosa a la base de la lengua y puede separarse cuando la cavidad se ensancha, como ocurre en las vocales /i, u/. La hendidura en forma de v de la lengua se observa muy marcada en la parte superior de los cortes coronales 6 y 7 en la vocal /i/. En estos cortes, se pueden observar los

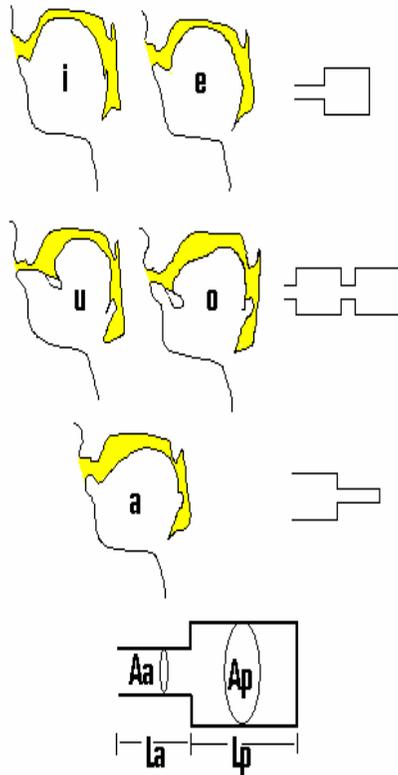
dientes inferiores (en negro) delineados por las encías (gris claro). En el corte 7 de la vocal /i/ puede observarse la úvula contra la pared posterior de la faringe (parte inferior de la sección de área).

En los cortes transversales la hendidura lingual se observa en la parte inferior de los cortes 10,11,12,13 en la vocal /a/, en el 10,11,12 para /i/ y 10,11 para /u/. En esta zona se observa el arco del paladar más pronunciado. La porción inferior de las secciones de área de las cavidades aéreas en los cortes 13, 14, y 15 son difíciles de delimitar por la falta de contrastes adecuados y por la interferencia producida por los otros cortes simultáneos, en la forma de barras negras transversales. Para resolver esta indefinición se ha recurrido a los cortes sagitales que permiten determinar la altura media de la sección de área.

Las áreas se han dibujado a mano alzada, calcando el contorno en una visión ampliada por tres, sobre la pantalla. La indefinición de los contornos produce variaciones en la medición influidas por la decisión del operador frente a su necesidad de mantener la consistencia con los contornos adyacentes, el conocimiento a priori que este tiene de la anatomía del tracto y del aumento o disminución de la superficie esperada en cada sección. También influye el trazado, el grado de contraste y brillo con la que se ajusta la imagen. A pesar de todos estos factores las mediciones reflejan el comportamiento general esperado en cada articulación vocal.

Las áreas y longitudes medidas, fueron utilizadas para calcular las frecuencias de formantes. Los cálculos se efectuaron con las expresiones derivadas de la

aplicación de la solución de la ecuación de onda en tubos de sección uniforme (Fant, 1960).



$$F_n = \frac{C(2n+1)}{4L}$$

para tubos abiertos en un extremo y cerrado en el otro extremo.

$$F_n = \frac{C(n+1)}{2L}$$

para tubos abiertos o cerrados en ambos extremos.

$$F_H = \frac{C}{2\pi} \sqrt{\frac{Aa}{LpApLa}} \quad (3)$$

para tubos compuestos. (Resonador de Hemholtz).

Donde C: velocidad del sonido en el aire, L: longitud del tubo, A: la sección de área,  $\pi = 3,14$ , a: anterior, p: posterior y  $n = 0,1,2,3...$

**Figura 10.** Esquemas de los cortes sagitales, configuración de tubos equivalente y ecuaciones utilizadas para el cálculo de los formantes para los tubos parciales constituyentes.

**Tabla II.** Longitudes y sección de área promedio para las vocales /i, a, u/ obtenidas de las mediciones por IRM.

Vocal	Lp	Ap	La	Aa	Lp	Ap	La	Aa	LTotal
/i/	7.9	4	6.1	0.89					14.50
/a/	9.2	-	6.1	-					15.30
/u/	4.5	3.47	5	0.59	4.5	4.5	1	0.53	15.00

La longitud del tracto vocal varía para cada vocal tal como se indica en L Total. La longitud total indicada en la Figura 9 (14 cm) corresponde a las distancias constantes entre los puntos fijos de medición de la sección de área.

**Tabla III.** Valores de formantes medidos acústicamente (en posición Supina (Su) y Sedente (Se) ) y valores de formantes calculados a partir de las áreas y longitudes medidas en las IRM. Se indican las diferencias con el valor calculado en semitonos (ST).

Vocal		F1	ST	F2	ST	F3	ST	F4	ST
/i/	Med. Su	406	2.17	2800	2.7	3290	2.6	4300	0.2
	Med Se	406	2.17	2780	2.6	3400	3.1	4300	0.2
	Calculado	358		2388		2819		4777	
/u/	Med Su	397	3.0	773	2.2	-		-	
	Med Se	418	2.1	825	1.	-		-	
	Calculado	472		875		-		-	
/a/	Med Su	790	2.89	1300	0.4	3000	1.3	4260	1.1
	Med Se	932	0.03	1320	0.1	3060	1.7	4000	0.04
	Calculado	934		1400		2778		3990	

Las grabaciones de las vocales se realizaron a posteriori de la obtención de las IRM, en dos posiciones: supina y sedente. Como parámetros de muestreo se emplearon una frecuencia de 16kHz y un nivel de muestreo de 16 bits. Las mediciones acústicas de los formantes se realizaron, con el sistema de análisis de sonidos de habla ANAGRAF (Gurlekian y otros, 1994) utilizando el método de predicción lineal (LPC = 14) obtenido por auto correlación. Los valores medidos se obtuvieron en el instante de máxima energía total.

## VI. CONCLUSIONES

Las imágenes articulatorias realizadas en el tracto vocal durante la emisión de vocales sostenidas mediante la técnica de IRM presentan una definición de alta calidad. Esta calidad permite calcular los formantes vocálicos con mayor precisión que las técnicas empleadas anteriormente. Las mediciones acústicas de los formantes en posición supina y sedente, realizadas en otra sesión, reflejan valores semejantes a los obtenidos por las ecuaciones que utilizan los parámetros articulatorios medidos.

## VII. AGRADECIMIENTOS

Los autores desean agradecer a la fundación Jaime Rocca por la ayuda brindada en el empleo del equipo de IRM del Hospital Italiano y a la Dra. Dora Latourrete por su colaboración en la realización de este trabajo.

## RESUMEN

En este trabajo se realiza la descripción articulatoria de vocales -mediante la técnica de adquisición de imágenes por resonancia magnética (IRM) del tracto vocal de un sujeto femenino normal, hablante del español de Buenos Aires. La medición de longitudes en los cortes sagitales, y de las áreas del tracto vocal en los cortes coronales, oblicuos y axiales, se realizó para las cinco vocales emitidas en forma aislada y sostenida. Las mediciones articulatorias efectuadas en vocales permitieron estimar la frecuencia de los formantes vocálicos. Estos valores fueron comparados con los producidos acústicamente por el mismo sujeto al emitir las vocales. Las diferencias observadas no son significativas y se discuten las posibles causas que afectan la medición.

## ABSTRACT

This paper presents the articulatory description of vowels by means of the MRI technique. They were measured in a normal female vocal tract speaker of Buenos Aires Argentine Spanish. Sagittal cuts provided length measurements, while coronal, oblique and axial cuts allowed the measurements of area sections of the five vowels, uttered sustained in isolation. Articulator measurements of vowels allowed the estimation of formant values. The predicted formant values were matched with their acoustic counterparts. Differences are small and the artefacts affecting the articulatory measurements are discussed.

## REFERENCIAS

**Apostol, L., Perrier, P. y Badin, P.** 3D geometry of the vocal tract and interspeaker variability. *Proc ICPhS 99*, 443-446.

**Baer T., Gore J. C., Gracco L. C. y Nye P. W.** Analysis of vocal tract shape and dimension using magnetic resonance imaging: vowels. *J. Acoust. Soc. Am.* 90 (2), 799-828, 1991.

**Beautemps,, D., Badin, P. y Laboissiere, R.** Deriving vocal tract area functions from midsagittal profiles and formant frequencies: a new model for vowels and fricative consonants based on experimental data. *Speech Commun.* 16, 27-47, 1995.

**Chiba, T. y Kajiyama, M.** The vowel, its nature and structure. *Tokio-Kaiseikan, Tokio, 1941*

**Dang J y Honda K.** Morphological and acoustical analysis of the nasal and paranasal cavities. *J. Acoust. Soc. Am.* 96, 2088-2100, 1994.

**Demolin D., Metens T. y Soquet A..** Three dimensional measurements of vocal tract by MRI. *Proc ICSLP 96*, 1: 272-275.

**Engwall, O. y Badin P.** Collecting and analysing two- and three- dimensional MRI data from Swedish. *KTH TMH-QPSR 3-4, 1999.*

**Fant, G.** The acoustic theory of speech production. *Moulon, The Hague, 1960.*

**Fitch T. W y Giedd J.** Morphology and development of the human vocal tract: a study using magnetic resonance imaging. *J. Acoust. Soc. Am.* 106(3), 1511-1522, 1999.

**Heinz J.M. y Stevens K.N.** On the derivation of area functions and acoustic spectra from cineradiographic films of speech. *J. Acoust. Soc. Am.* 36, 1037(abs), 1964.

**Hoole P y Nguyen N.** Electromagnetic Articulography in coarticulation research. *Forschungsberichte des Instituts für Phonetik und Spachliche Kommunikation der Universität München, FIPKM*, 35:177-184, 1997.

**Maeda S.** On articulatory and acoustic variabilities. *Journal of Phonetics*, 19:321-331, 1991.

**Mair S.J., Scully C.y Shadle C.H.** Distinctions between [t] and [tsh] using electropalatography data. *Proc ICSLP 96*, 1597-1600, 1996.

**Matsumura M., Niikawa T., Shimizu K. y Hashimoto Y.** Measurement and analysis of 3D shapes of vocal tract, dental crown and nasal cavity using MRI: vowels and fricatives. *Proc ICSLP 94, S12: 13.1-13.4, 1994.*

- Mermelstein, P.** Articulatory model of speech production. *J. Acoust. Soc. Am.* 53, 1070-1082, 1973.
- Moore, C. A.** The correspondence of vocal tract resonance with volumens obtained from magnetic resonances images. *J. Speech Hear. Res.* 35, 1009-1023, 1992.
- Narayanan, S. S., Alwan, A. A y Haker, K.** An articulatory study of fricative consonants using magnetic resonance imaging. *J. Acoust. Soc. Am.* 98, 1325-1347, 1995.
- Narayanan, S. S., Alwan, A. A y Haker, K.** Towards articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part I. The laterals. *J. Acoust. Soc. Am.* 101, 1064-1077, 1997.
- Narayanan, S. S., Alwan, A. A y Haker, K.** Towards articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part I. The rothics. *J. Acoust. Soc. Am.* 101, 1077-1089, 1997.
- Navarro Tomás, T.** Manual de pronunciación española. Madrid: Gredos, 1918.
- Nguyen N y Marchal A.** Assessment of an electromagnetometric system for the investigation of articulatory movements in speech production. *J. Acoust. Soc. Am.*, 91(2):1152-1155. EMA., 1992.
- Perkell, J. S.** Physiology of speech production: results and implications of a quantitative cineradiographic study. MIT, Cambridge, MA, 1969.
- Perkell J.S., Cohen M.H., Svirsky M.A., Matthies M.L., Garabieta I. y Jackson M.T.T.** Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements. *J. Acoust. Soc. Am.*, 92:3078-3096, 1992.
- Perrier, P., Boe, L. J. y Sock, R.** Vocal tract area function from midsagittal dimensions with CT scans and vocal tract cast: modelling the transition with two sets of coefficients. *J. Speech Hear. Res.* 35, 53-67, 1992.
- Quilis, A.** Tratado de fonología y fonética españolas. Madrid: Gredos (Biblioteca Romántica Hispánica, manuales, 74).
- Schönle P.W.** The developmental genealogy of Electromagnetic Articulography (EMA). *Forschungsberichte des Instituts für Phonetik und Spachliche Kommunikation der Universität München, FIPKM*, 30:83-90, 1992.
- Soquet A., Lecuit V., Metens T., and Demolin D.** From sagittal cut to area function: an MRI investigation. *Proc ICSLP 96*, 1205-1208, 1996.
- Stone, M.** A three dimensional model of tongue movement based on ultrasound and x-ray microbeam data. *J. Acoust. Soc. Am.* 87, 2207-2217, 1990.
- Story, B.H., Titze, I. R y Hoffman, E. A.** Vocal tract area function from magnetic resonance imaging. *J. Acoust. Soc. Am.* 100, 537-554, 1996.
- Sundberg, J.** On the problem of obtaining area functions from lateral x-ray pictures of the vocal tract. *KTH, STL-QPSR 1: 43-45, 1969.*
- Toledo, G. A y García Jurado, M. A.** Análisis radiográfico de vocales españolas. *II Congreso Nacional de Lingüística, San Juan, Argentina, 1981.*
- Yang C-S y Kasuya H.** Accurate measurement of vocal tract shapes from magnetic resonance images of child, female and male subjects. *Proc ICSLP 94*, 623-626, Yokohama, Japan.

**Ing. Jorge A. Gurlekian**  
Investigador del CONICET  
Laboratorio de Investigaciones Sensoriales  
Instituto de Neurociencias Aplicadas. Hospital de Clínicas  
Av. Córdoba 2351, 9 Piso Sala 2  
TE: 5950-9024  
[jag@fmed.uba.ar](mailto:jag@fmed.uba.ar)

**Fga. Natalia Elisei**  
Fonoaudióloga residente del Hospital de Clínicas  
En rotación en el Laboratorio de Investigaciones Sensoriales

**Dr. Martín Eleta**  
Médico de planta del Servicio de Diagnóstico por Imágenes del Hospital Italiano de Buenos Aires.