

# Implementación de una base de datos de alteraciones de la voz y el habla del español de Buenos Aires

Dra. Natalia Gabriela Elisei (1), Médica ORL Dora Latourrette (2), Lig. Fga. María Eugenia Pérez Ibañez (3), Lic. Santiago Giusti (4)

## Resumen

La descripción acústica y perceptual de patologías de la voz y el habla se observa en la mayoría de los trabajos presentados como soporte de las distintas estrategias de diagnóstico y tratamiento. Las descripciones cualitativas y cuantitativas consideradas como referencia normal no son aplicables a todas las lenguas por igual, no sólo por tener diferencias esenciales en el sistema fonológico, semántico y gramatical, sino por las variaciones fonéticas, de entonación y de ritmo propias de una misma lengua.

En la mayoría de los trabajos clínicos y científicos de esta especialidad o vinculados a las patologías de la voz, se utiliza una base de datos comercial americana. En nuestro país han existido intentos de proyectos frustrados. El objetivo de este trabajo fue contribuir con el desarrollo de una base de datos inédita, para el español de Buenos Aires, que contemple la diversidad de los desórdenes vocales.

Para el reclutamiento de los participantes se solicitó la colaboración del Servicio de Otorrinolaringología y Fonoaudiología del Hospital de Clínicas (HCJSM), dependiente de la Universidad de Buenos Aires; del Servicio de Otorrinolaringología y Fonoaudiología del Hospital Italiano de Buenos Aires (HI) y del Instituto Superior de Enseñanza Radiofónica (ISER). La creación de esta base de datos involucró cuatro etapas: diseño, grabación, edición e implementación.

En la etapa de diseño y grabación se registraron los hablantes del español de Buenos Aires, de ambos sexos, con edades comprendidas entre los 17 y los 80 años, con y sin desórdenes vocales diagnosticados por fibrolaringostroboscopia. Se excluyeron pacientes con problemas auditivos. En la etapa de edición se realizaron los ajustes necesarios a los archivos de sonido para su análisis acústico posterior.

Para poder modelar, organizar y administrar la información de manera dinámica y eficiente, se de-

finió un modelo relacional (base de datos relacional) implementado en Microsoft Access 2007.

En este modelo todos los datos son almacenados en relaciones, y como cada relación es un conjunto de datos, el orden en el que estos se almacenan no posee relevancia. Esto tiene la considerable ventaja de que es más fácil de entender y de utilizar por un usuario no experto. La información puede ser recuperada o almacenada por medio de consultas que ofrecen una amplia flexibilidad y poder para administrar la información.

En este trabajo se han presentado el diseño y la construcción de la base de datos utilizada para implementar distintos tipos de investigaciones a posteriori. Se ha descrito la organización del corpus teniendo en cuenta las variables clásicas que pueden influenciar el material vocal, como son: sexo, edad, grupo (normal o patológico) y diagnóstico médico.

Este trabajo contribuye a generar una base de datos pasible de ser utilizada tanto por investigadores y profesionales afines al área así como graduados, residentes y alumnos.

**Palabras claves:** Base de datos – multimedia - voz normal y patológica - español de Buenos Aires.

## Introducción

La descripción fisiológica, acústica y perceptual de patologías de la voz y el habla se observa en la mayoría de los trabajos presentados como soporte de las distintas estrategias de diagnóstico y tratamiento. Las descripciones cualitativas y cuantitativas en el área de la voz consideradas de referencia normal no son aplicables a todas las lenguas por igual, no sólo por tener diferencias esenciales en el sistema fonológico, semántico y gramatical, sino por las variaciones fonéticas, de entonación y ritmo propias de una misma lengua.

Es por esto que surgió la necesidad de generar una base de datos que permita asociar los desórde-

1- Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) E-mail: natalia.elisei@gmail.com. Tel.: 4703-1391.

2- Hospital Italiano de Buenos Aires.

3- Instituto Superior de Enseñanza Radiofónica.

4- Ciencias de la Computación.

nes de la voz y el habla a otras variables del análisis fisiológico, acústico y perceptual.

Numerosos estudios se valieron de métodos instrumentales como la estroboscopia laríngea y el análisis acústico como "gold-standard" para establecer una comparación y así acceder al valor pronóstico de medidas acústicas para evaluación objetiva de la voz (1-3). Se investigaron alteraciones en la voz (4-9), influencia del entrenamiento vocal (10, 11), comparación de lenguas (12), diferencias en intensidad vocal (13), diferencias de sexo (14-16) e identificación de hablantes (17).

La dificultad que se plantea en el uso de un solo método para evaluar los desórdenes vocales radica, fundamentalmente, en que son muy diferentes los procesos patológicos que afectan a varios aspectos del desempeño vocal y en diferentes grados. Es más, tanto las mediciones acústicas como la estroboscopia individualmente brindan solo una parte de la información contenida, por ejemplo, en el análisis perceptual. Por esta razón muchos equipos han propuesto modelos de análisis multivariados para realzar el alcance de los datos (18-24).

Desde entonces, el tema de investigación más relevante ha sido diseñar un protocolo multiparamétrico usando varios parámetros acústicos, aerodinámicos y de diagnóstico por imágenes, como la estroboscopia laríngea, para la evaluación de los desórdenes vocales.

En la mayoría de los trabajos clínicos y científicos de esta especialidad o vinculados a las patologías de la voz se utiliza una base de datos comercial americana para testear los protocolos multiparamétricos anteriormente mencionados. Kay Elemetrics (25) ha registrado una base de datos que desarrolló el Laboratorio de Voz y Habla de Massachusetts (MEEI). Contiene 1.400 muestras de voz de aproximadamente 700 sujetos, incluyendo la fonación de vocales sostenidas y muestras de habla conectada de pacientes con una amplia variedad de desórdenes vocales orgánicos, neurológicos, traumáticos y psicógenos, así como de 53 normales. Todos los pacientes presentan una evaluación vocal que incluye estroboscopia, mediciones aerodinámicas, fonatorias, y análisis acústico. La información clínica contiene el diagnóstico junto con datos vinculados al paciente, como: edad, sexo, si es fumador o no, etc. En Francia, el Laboratorio de Clínica Fonoaudiológica en el Hospital Universitario Timone, en Marsella, en conjunto con el Laboratorio Parole et Language d'Université de Provence (26), han desarrollado una estación de trabajo EVA® con su respectiva base de datos de 449 hablantes francófonos (391 casos con patología y 58 normales).

En nuestro país han existido intentos de proyectos frustrados que no han logrado concretar el objetivo de generar una base de datos propia del español de Buenos Aires, básicamente debido a la falta de presupuesto y tiempo en los ámbitos clínicos hospitalarios.

En la presente investigación se asumió un concepto amplio de la voz, como un fenómeno multiparamétrico, y se conjugaron los aportes de varias disciplinas (otorrinolaringología, fonoaudiología, informática, lingüística y física acústica, etc.) con el objetivo de contribuir al desarrollo de una base de datos multimedia e inédita, a nuestro entender, para el español de Buenos Aires.

La creación de esta base de datos involucró cuatro etapas: diseño, grabación, edición e implementación que se describen a continuación.

## Descripción de la base de datos

### I. Diseño

Para el reclutamiento de los participantes se solicitó la colaboración al EBAP1, otra institución asociada EBAP2 y EBAN, para realizar la toma de registros. Las dos instituciones médicas EBAP1 y EBAP2 se eligieron por contar con equipamiento pertinente para el diagnóstico de patologías vocales (estroboscopia) y por el amplio espectro de consultas que reciben a diario. El EBAN se eligió por concentrar voces ya seleccionadas como patrones de alto rendimiento vocal en voz hablada. El protocolo de trabajo fue aprobado por los Comités de Ética de sendas instituciones.

Dentro del universo de todos los sujetos con y sin desórdenes vocales, nuestra población accesible fueron aquellos pacientes que concurrieron a las instituciones médicas EBAP1 y EBAP2 desde la fecha de aprobación del protocolo hasta diciembre de 2010 y alumnos del EBAN cursando sus estudios durante 2009. Los criterios de inclusión y exclusión se detallan a continuación:

**Criterios de inclusión:** Hablantes del español de Buenos Aires; de ambos sexos con edades entre 17 y 80 años; con y sin desórdenes vocales; con diagnóstico estroboscópico.

**Criterios de exclusión:** Pacientes con problemas auditivos.

Finalmente, la muestra a analizar en este trabajo quedó conformada por los sujetos que concurrieron a estas instituciones durante los plazos descriptos, con y sin desórdenes vocales, que consintieron en participar del estudio. Todos los sujetos aceptaron libremente y de forma espontánea participar conociendo que las muestras de voz serían analizadas

y, finalmente, presentadas al público, pero que sus identidades no serían reveladas.<sup>1</sup>

### Tareas

Las características acústicas de la voz como la duración, la energía y la frecuencia fundamental de vibración, junto con la características espectrales determinadas por lo formantes, son los parámetros básicos que en sus variaciones se correlacionan con una diversidad de alteraciones vocales. Por ello, se registraron las emisiones de vocales aisladas y un conjunto de oraciones fonéticamente balanceadas siguiendo los criterios expuestos en (27). Para el diseño de este corpus se utilizó un transcriptor automático del alfabeto fonético que permitió la creación de dos párrafos conteniendo todos los fonemas y los principales alófonos utilizados en el país.

También se registraron digitalmente las imágenes estroboscópicas de los sujetos evaluados. La estroboscopia constituye el método más valioso de exploración laríngea; es el elemento más importante de los laboratorios de voz. La apertura y el cierre de las cuerdas vocales ocurren demasiado rápido y la percepción visual humana solo puede detectar objetos en movimiento periódico cuando su frecuencia es menor de 6 ciclos/s; todo movimiento que exceda esta frecuencia es percibido como estático. La estroboscopia permite transformar lo que ocurre 100 o más veces por segundo en una imagen que dura 1 segundo, haciéndolo fácilmente perceptible a la vista humana.

### Instrucciones

En la estroboscopia con endoscopio rígido el paciente permanece sentado sobre un sillón; la posición de la cabeza y de la barbilla dependerá del ángulo que presente el endoscopio elegido. La introducción del endoscopio se realiza a través de la boca y con frecuencia el explorador sujeta la lengua del paciente con ayuda de una gasa estéril para facilitar su colocación. Puede requerir el uso de un anestésico previo a la introducción del endoscopio si el paciente presenta náuseas durante el estudio.

En una evaluación posterior en sala sonoamortiguada, se les solicitó a los participantes que pronunciaran en tres oportunidades las vocales del español de Buenos Aires de manera sostenida (tiempo estimado de 3 a 5 segundos), a una intensidad y frecuencia espontáneas. Luego se pidió que leyeran los enunciados presentados, una cada vez como si estuviera hablando con alguien en una conversación

real. Se utilizó para esto el programa GRABAR desarrollado en el LAB 1 como interfase, que permite a través de la presentación de sucesivas pantallas, la presentación y grabación de las distintas muestras vocales del protocolo.

### I. Grabación

Previo al registro de sus voces, los participantes respondieron a un breve cuestionario relacionado con factores de riesgo. Luego se registraron las vocales y las frases de habla conectada. Este material fue emitido por los hablantes a una intensidad de habla conversacional y registrado en forma digital con un nivel de calidad de 16 bits y una frecuencia de muestreo de 44.100 muestras por segundo y no se utilizó ningún tipo de compresión.

Para la captura digital de las imágenes se utilizó el sistema ENDODIGI - PAM Argentina SA. Los diagnósticos etiológicos presentes en la base de datos son: lesiones estructurales mínimas, congestión por reflujo gastroesofágico, papilomatosis, granulomas, hiperfunción, hiperplasia, queratosis, edema de cuerdas vocales, pólipos cordales, fonación ventricular, tejido de cicatrización, temblor vocal, estenosis laríngeas y parálisis cordales, entre otras.

### Equipamiento

Para realizar los estudios fisiológicos se utilizó un estroboscopio Ecleris que cuenta con un sistema electrónico de control de luz (Xenon Auto Transition Technology) que permite funcionar con la misma lámpara de xenón de 180 watts, tanto en modo continuo como estroboscópico. Se utilizó una fibra óptica rígida de 0° o fibroscopio flexible Olympus en caso de que no toleraran la fibra óptica rígida.

Como los registros se realizaron en tres instituciones diferentes, se detallan los equipos que se utilizaron. En el EBAN las emisiones fueron grabadas digitalmente en una computadora portátil, utilizando una placa de sonido externa USB marca Andrea Electronics modelo Pureaudio. Se utilizó un micrófono AGK D880, tipo dinámico unidireccional cardioide, con un rango de frecuencia de 60 Hz-20 kHz, sensibilidad de 2,5 mV/Pa (-52 dBV) e impedancia de 600 Ohm situado a 10 cm de la boca en una sala sonoamortiguada.

En el EBAP1 las emisiones fueron grabadas digitalmente en una computadora de escritorio utilizando una placa de sonido externa USB marca M-Audio Firewire modelo 1410. Se utilizó un micrófono AGK D770, tipo dinámico unidireccional cardioide, con un rango de frecuencia de 60 Hz-20 kHz, sensibilidad de 2,5 mV/Pa (-52 dBV) e impedancia de 600 Ohm situado a 10 cm de la boca en una sala acústica

<sup>1</sup> Ley Nacional de Protección de Datos Personales 25.326 (Ley de Habeas Data) [58], en concordancia con la normativa internacional sobre registro de enfermedades y protección de datos personales y privados, de acuerdo con 18ª Asamblea Médica Mundial de Helsinki (1964).

y antecámara con nivel de ruido de 35 dB y tiempo de reverberación menor a 1 segundo.

En el EBAP2 las emisiones fueron grabadas digitalmente en una computadora de escritorio utilizando una placa de sonido externa USB marca Creative Labs modelo SB1090. Se utilizó un micrófono SHURE 8900, tipo dinámico unidireccional cardioide, con un rango de frecuencia de 50 Hz-15 kHz, sensibilidad de 2,5 mV/Pa (-52 dBV) e impedancia de 600 Ohm situado a 10 cm de la boca en una sala sonoamortiguada.

## II. Edición de las muestras

Una vez que se tomaron las muestras se editaron a través del programa Sound Forge Versión 8.0b. La edición de cada sonido se realizó tomando el cuerpo de la emisión y desechando el ataque y la filatura de cada muestra. Una vez realizado esto, se normalizó por el método RMS para poder comparar los archivos. Cada vocal tiene una duración que va de 1 a 5 segundos. Los párrafos fueron editados en una sola muestra, a la que se le extrajeron automáticamente, a través de la función de PRAAT [28] LTAS pitch-corrected, los sonidos sordos y las pausas. Estas muestras "corregidas", conteniendo solo los sonidos sonoros, se utilizaron para su posterior análisis espectral. Las imágenes de video no fueron editadas.

## III. Implementación

Para poder modelar, organizar y administrar la información recogida de manera dinámica y eficiente, se definió un modelo relacional (base de datos relacional) implementado en Microsoft Access 2007. En este modelo todos los datos son almacenados en relaciones, y como cada relación es un conjunto de datos, el orden en el que estos se almacenan no posee relevancia. Esto tiene la considerable ventaja de que es más fácil de entender y de utilizar por un usuario no experto. La información puede ser recuperada o almacenada por medio de consultas que ofrecen una amplia flexibilidad y poder para administrar la información.

Una base de datos relacional es un conjunto de una o más tablas estructuradas en registros (líneas) y campos (columnas), que se vinculan entre sí por un campo en común, en ambos casos posee las mismas características como por ejemplo el nombre de campo, tipo y longitud; a este campo generalmente se le denomina ID, identificador o clave. Este modelo considera la base de datos como una colección de relaciones. De manera simple, una relación representa una tabla que no es más que un conjunto de filas, cada fila es un conjunto de campos y cada campo representa un valor que, interpretado, describe el mundo real. Cada fila también se puede

denominar como tupla o registro y a cada columna también se la puede llamar campo o atributo. Una clave primaria es una clave única, elegida entre todas las candidatas, que define unívocamente a todos los demás atributos de la tabla, para especificar los datos que serán relacionados con las demás tablas. La forma de hacer esto es por medio de claves foráneas. Sólo puede existir una clave primaria por tabla y ningún campo de dicha clave puede contener valores nulos.

Una clave foránea es una referencia a una clave en otra tabla. Las claves foráneas no necesitan ser claves únicas en la tabla donde están y sí a donde están referenciadas. El modelo de datos consta de las siguientes entidades y los nombres de tablas respectivas: pacientes (PATIENTS), diagnósticos (DIAGNOSTICS), fuentes (SOURCES), protocolos (PROTOCOLS), ondas de sonido (WAVES), videos de laringoestroboscopia (VIDEOS) y tablas secundarias con las sucesivas mediciones acústicas. El modelo de datos anteriormente descrito puede resumirse gráficamente en el diagrama de entidad-relación de la Figura 1.

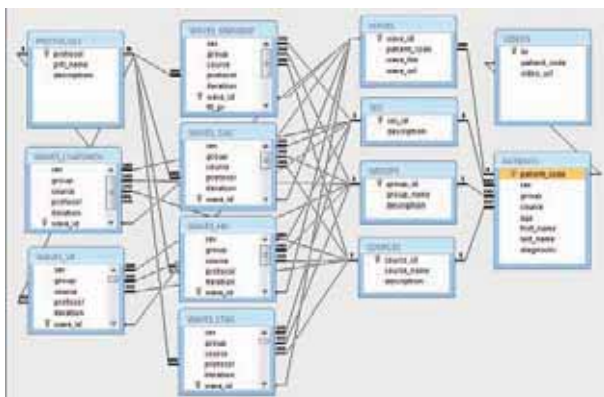


Figura 1. Diagrama de entidad-relación de la base de datos de voces.

Para facilitar la visualización y la carga de los datos también se diseñaron pantallas o formularios. De esta manera, posicionado en un sujeto específico, se puede acceder a todas las métricas obtenidas con los distintos métodos y procedimientos, así como también a las distintas grabaciones que se efectuaron (Figura 2).

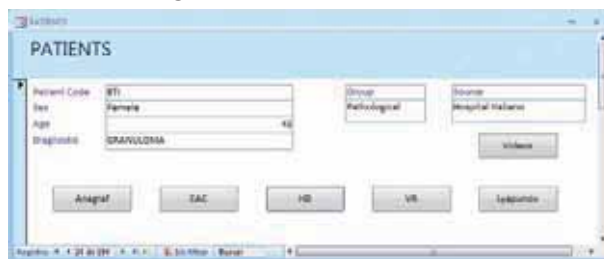


Figura 2. Interface de usuario de la base de datos.

Para complementar la información de la persona, en los casos que fue posible concretar el estudio, se dispone de estroboscopías digitalizadas para analizar y contrastar con el resto de la información (Figura 3).



Figura 3. Imágenes de videoestroboscopías laríngeas.

### Estadísticas descriptivas de la base de datos implementada

La base está integrada por un total de 194 sujetos hablantes del español de Buenos Aires, de los cuales 78 (40,2%) fueron hombres y 116 (59,8%) mujeres.

La media de edad fue de 36,35 años y con una desviación estándar de 16,06. Las voces de individuos normales pertenecen a 33 hombres y 33 mujeres, con edades promedio de 27,38 ± 7,9 y 26,78 ± 7,9 años respectivamente (se indica el valor medio y ± el desvío estándar).

El conjunto de voces patológicas contiene muestras de 45 hablantes masculinos y 83 femeninos. Las edades promedio en este caso son de 45,88 ± 22,02 y 38,31 ± 15,68 años para el grupo de hablantes masculinos y femeninos, respectivamente. La distribución de los datos según la procedencia o fuente se observan en la siguiente Tabla.

Fuente	N	Media	SD	Rango	Mediana
EBAP1	50	44.22	21.893	62	46
EBAP2	78	38.91	12.881	56	40
EBAN	66	27.08	7.852	33	24.5
<b>Total</b>	<b>194</b>	<b>36.35</b>	<b>16.059</b>	<b>62</b>	<b>32</b>

Tabla 1. Medias y desviaciones de la muestra según las edades y la procedencia. Referencias: EBAP1, EBAP2 y EBAN.

Los diagnósticos etiológicos presentes en la base de datos son: lesiones estructurales mínimas, congestión por reflujo gastroesofágico, papilomatosis, granulomas, hiperfunción, hiperplasia, queratosis, edema de cuerdas vocales, pólipos cordales, fonación ventricular, tejido de cicatrización, temblor vocal, estenosis laríngeas y parálisis cordales, entre otras.

La diversidad en la nomenclatura de los diagnósticos etiológicos hacía prácticamente imposible la comparación de los datos. Por este motivo se recurrió a la codificación de los diagnósticos según el Diagnostic Classification System for Voice Disorders (DCSV) (29).

Este sistema de clasificación de desórdenes vocales es un sistema "sindrómico". Es decir, incluye la terminología, las definiciones y los síntomas más probables, capturando lo esencial de la naturaleza de la enfermedad. El DCSV distingue a los no orgánicos de los orgánicos. Los trastornos de la voz

pueden distinguirse entre aquellos que tienen un origen orgánico (Trastornos Orgánicos de la Voz, TOV) y los llamados funcionales (Trastornos Funcionales de la Voz, TFV). Estos últimos, a su vez, se subdividen en dos grupos: trastornos psicogénicos de la voz (TPV) y trastornos de tensión muscular de la voz (TTMV). Los trastornos orgánicos se refieren a disfonías o afonías (ausencia total de voz) debidas a lesiones, cambios estructurales de las pliegues vocales o de las estructuras cartilaginosas, o interrupción de las innervaciones neurológicas de los mecanismos laríngeos. Son TOV generados por cambios tisulares, por ejemplo, los causados por una infección aguda, quistes, edemas, etc. También son ejemplos de TOV los traumas en la laringe que pueden ser ocasionados por heridas penetrantes o la fractura del cartílago tiroideo, y problemas neurológicos que causan las disfonías espasmódicas. Los trastornos funcionales son disfonías o afonías donde no existe una patología orgánica o, si la hay, es insuficiente para ser considerada responsable de la naturaleza y severidad del trastorno o es secundaria al problema funcional. Como se mencionó anteriormente, se pueden subdividir en: i) Trastornos psicogénicos de la voz: ocurre afonía o disfonía como resultado de un trastorno psicológico donde hay una pérdida súbita o intermitente del control voluntario sobre el inicio y mantenimiento de la fonación, en ausencia de patologías estructurales o neurológicas. Existen factores psico-sociales ligados a su causa; ii) Trastornos de tensión muscular de la voz: son disfonías que se desarrollan de manera gradual, como resultado de un proceso psicológico que lleva a patrones crónicos de mal uso de la musculatura laríngea. Con el tiempo, estos comportamientos pueden llevar al desarrollo de cambios orgánicos secundarios, como nódulos vocales. Un ejemplo es la disfonía por tensión muscular.

De la muestra codificada según DCSV el mayor porcentaje de diagnósticos, 81,25%, se contabiliza en la categoría de patologías orgánicas para ambos sexos, siendo más prevalente en el sexo femenino. En segundo lugar se ubican los trastornos por tensión muscular: 10,16%, y en tercer lugar aquellas patologías no especificadas, con un 7,03%. Ver Tabla 2.

Categoría diagnóstica	EBAP1		EBAP2		Total	Porcentaje
	F	M	F	M		
TTMV	2	1	6	4	13	10.16
TOV	24	21	45	14	104	81.25
TPV	1	1	0	0	2	1.56
MISC	0	0	5	4	9	7.03
<b>Total</b>	<b>27</b>	<b>23</b>	<b>56</b>	<b>22</b>	<b>128</b>	<b>100</b>

Tabla 2. Distribución por categoría diagnóstica según DCSV para ambas instituciones (EBAP1 y EBAP2). Referencias: TTMV (Trastornos por tensión muscular vocal), TOV (Trastornos orgánicos de la voz), TPV (Trastornos psicogénicos de la voz), MISC (Misceláneos o no especificados).

En el proceso de grabación se cuidó que las señas no tuviesen un nivel de intensidad inadecuado y que las distintas vocales tuviesen una duración de, al menos, un segundo; que los ambientes donde se realizaron las grabaciones fueran habitaciones con los niveles de contaminación acústica controlados, intentando obtener un compromiso entre muy bajos niveles de ruido y niveles de ruido realistas, para que los datos medidos sobre esta base resultaran comparables a otros datos registrados en las condiciones habituales de trabajo de cualquier ámbito clínico u hospitalario.

Se consideró que el factor de procedencia o la fuente podía interactuar como variable confundente. Por tal motivo, se estudiaron los 3 grupos y se obtuvieron diferencias significativas entre EBAP1 y EBAP2. Se presenta a continuación una descripción detallada de los datos desagregada por la fuente del dato. Ver Figura 3.

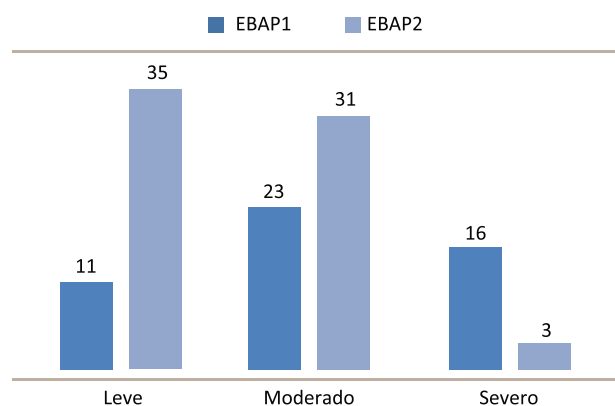


Figura 3. Distribución por niveles de severidad para ambas instituciones (EBAP1 y EBAP2).

Dado que en ambos grupos prevalece la categoría diagnóstica de TOV, se decide utilizar el score que ofrece el DCSVD en diagnósticos leves, moderados y severos, según sea el grado de severidad de la patología. En la siguiente figura se muestra el resultado de este análisis cualitativo. La cantidad de casos con diagnósticos moderados son semejantes en ambas fuentes. Sin embargo, el EBAP2 parece concentrar la mayor cantidad de casos leves y el EBAP1 capitalizar la mayor cantidad de casos severos. Se presume, entonces, que las diferencias en las variancias no están dadas por las condiciones diferentes de grabación, sino más bien a esta diferencia en los grados de severidad de los diagnósticos etiológicos.

A pesar de esto, se realizó también un estudio comparativo de 6 muestras tomadas al azar, 2 de cada fuente, una masculina y otra femenina. Se evaluó la relación señal a ruido y se verificó que el

ruido de fondo en todas las situaciones fuese semejante. Se considera que son las diferencias diagnósticas que influyen en las variancias de los grupos desagregados por el factor fuente y no las condiciones de grabación ni ambiente.

### Conclusiones

En este trabajo se ha presentado el diseño y la construcción de una base de datos multimedia inédita para el español de Buenos Aires. Se ha descrito la organización del corpus teniendo en cuenta las variables clásicas que pueden influenciar el material vocal, como son: sexo, edad, grupo (normal o patológico) y diagnóstico médico.

La base de datos contiene grabaciones de voz tanto normal como patológica, el registro digitalizado de las imágenes estroboscópicas del comportamiento laríngeo y los datos de las mediciones de los análisis acústicos correspondientes a los parámetros de la voz y a los rasgos distintivos del habla. Acompañando la información gráfica y sonora, se presenta información en forma de texto con datos del paciente, su historia clínica resumida y las características observadas. Los diagnósticos etiológicos presentes en la base de datos son: lesiones estructurales mínimas, congestión por reflujo gastroesofágico, papilomatosis, granulomas, hiperfunción, hiperplasia, queratosis, edema de cuerdas vocales, pólipos cordales, fonación ventricular, tejido de cicatrización, temblor vocal, estenosis laríngeas y parálisis cordales, entre otras.

Algunas ventajas de tener una base de datos sistematizada son que, por primera vez, los hospitales y las universidades pueden tener de manera rápida y de fácil acceso (en conveniente formato CD rom y, en un futuro próximo, online) a mucha información y muy rica en su contenido, con datos recolectados a través de un protocolo en diferentes instituciones, con diferentes grados de compromiso vocal grabados todos con idéntico protocolo y bajo condiciones similares.

También supone una ventaja para el entrenamiento perceptivo auditivo y visual, dado que los ejemplos y resultados de análisis pueden ser vistos y escuchados por el usuario en varias ocasiones y comparar resultados.

Esta base de datos ha demandado tiempo y esfuerzo, pero es un aporte válido luego de varios intentos frustrados. Si bien, inicialmente, fue utilizada para implementar diferentes tipos de mediciones acústicas, se espera que, luego de esta presentación, surjan nuevas propuestas de estudios y/o trabajos a realizar en conjunto con otras instituciones.

## Bibliografía

- [1] Nuñez Batalla F., Corte Santos P., Sequeiros Santiago G., Señaris González B. y Suárez Nieto C. *Evaluación perceptual de la disfonía: correlación con los parámetros acústicos y fiabilidad.* *Acta Otorrinolaringol Esp*, 55:282-287, 2004.
- [2] Dejonckere P.H., Remacle M., Fresnel-Elbaz E., Woisard V., Crevier-Buchman L., and Millet B. *Reliability and clinical relevance of perceptual evaluation of pathological voices.* *Rev Laryngol Otol Rhinol*, 117:247-258, 1998.
- [3] Låofqvist A. *The long time average spectrum as a tool in voice research.* *J Phonetics*, 14:471-476, 1986.
- [4] Fritzell B., Hallen O., and Sundberg J. *Evaluation of teflon injection procedures for paralytic dysphonia.* *Folia Phoniatica*, 26:414-421, 1974.
- [5] Prytz and Frokjaer-Jensen B. *Long-term average spectra analyses of normal and pathological voices.* *Folia Phoniatica*, 26:280-285, 1976.
- [6] Hammarberg B., Fritzell B., Gaußen J., Sundberg J., and Wedin L. *Perceptual and acoustic correlates of abnormal voice qualities.* *Acta Otolaryngol*, 90:441-451, 1980.
- [7] Hammarberg B. *Clinical Routines for the perceptual-acoustic assessment of dysphonia.* *Phoniatric and Logopedic Progress Report, Huddinge University Hospital, Sweden*, 1980; 4:14-29.
- [8] Kitzing P. and Akerlund L. *Long-time average spectrograms of dyphonic voices before and after therapy.* *Folia Phoniatica*, 45:53-61, 1993.
- [9] Hurme P. *Acoustic studies of voice variation.* *Jyväskylä studies in Communication 7.* *Jyväskylä: University of Jyväskylä*, 1996.
- [10] Wedin S., Leandersson L., and Wedin L. *Evaluation of voice training.* *Folia Phoniatica*, 30:103-112, 1978.
- [11] White B. *Singing techniques and vocal pedagogy.* *New York: Garland*, 1998.
- [12] Majewski W., Rothman H., and Hollien H. *Acoustic comparisons of american english and polish.* *J Phonetics*, 5:247-251, 1977.
- [13] Nordenberg M. and Sundberg J. *Effect on LTAS of vocal loudness variation.* *TMH-QPSR, KTH, Sweden*, 2003; 45:93-100.
- [14] Mendoza E., Valencia N., Muñoz J. y Trujillo H. *Differences in voice quality between men and women: Use of long-term average spectrum (LTAS).* *J Voice*, 10:59-66, 1996.
- [15] White P. *Long-term average spectrum (LTAS) analysis of sex- and gender-related differences in children's voices.* *Logopedics Phonetics Vocology*, 26:97-101, 2001.
- [16] White P. *The effect of vocal intensity variation on children's voices using long-term average spectrum (LTAS) analysis.* *Logopedics Phonetics Vocology*, 23:111-120, 1998.
- [17] Hollien H. *The acoustics of crime. The new science of forensic phonetics.* *Plenum: New York*, 1990.
- [18] Doscov, D. & Ivanov, T. "Computer Analysis of the voice of Patients with Organic Dysphonia". *Scandinavian Journal Of Logopedics and Phoniatics*, 1993, 18, 1, 29-32.
- [19] de Krom, G. *Some spectral correlates of pathological breathy and rough voice quality for different types of vowel fragments.* *J Speech Hear Res*, 38:794-811, 1995.
- [20] de Oliveira Rosa, M, Pereire, J.C y Grellet, M. *Adaptative estimation of residue signal for voice pathology diagnosis.* *IEEE Trans. Biomed. Eng.*, Vol. 47, Nº 1, 96-102, Jan. 2000.
- [21] Yanagihara, N. *Significance of harmonic changes and noise components in hoarseness.* *J Speech Hear Res*, vol. 10, no. 3, pp 531-541, 1967.
- [22] Kasuya, H, Mashima, K y Ebihara, S. *Normalized noise energy as an acoustic measure to evaluate pathologic voice.* *J. Acoust. Soc. Am.*, vol. 80, no. 5, pp. 1329-1334, 1986.
- [23] de Krom, G. *A cepstrum-based technique for determining a harmonic-to noise ratio in speech signals.* *J. Speech. Hear. Res*, vol. 36, pp. 254-266, 1993.
- [24] Michaelis, D., Frohlich, M. and Strube, H. W. *Selection and combination of acoustic features for the description of pathologic voices"* *J. Acoust. Soc. Am.*, vol. 103, no. 3, pp. 1628-1639, 1998.
- [25] <http://www.kayelemetrics.com/>
- [26] Ghio, A. et al. *Corpus de «parole pathologique». État d'avancement et enjeux méthodologiques au LPL, Travaux Interdisciplinaires du Laboratoire Parole et Langage*, vol. 25, p. 109-126, 2006.
- [27] Gurlekian, J. A., Colantoni, L., Torres, H., Rincón, A., Moreno, A. and Mariño, J. "Database for an Automatic Speech Recognition System for Argentine Spanish", *Proceedings of the IRCS Workshop of Linguistic Databases.* Bird, Buneman & Liberman (eds). Philadelphia, pp. 92-98, December 2001.
- [28] PRAAT "Doing phonetics by computer", versión 4.6.06 de uso liberado y público en [www.fon.hum.uva.nl/praat](http://www.fon.hum.uva.nl/praat)
- [29] Baker J, Ben-Tovim DI, Butcher A, Esterman A, McLaughlin K. *Development of a modified diagnostic classification system for voice disorders with inter-rater reliability study.* *Logoped Phoniatr Vo- col.* 2007;32(3):99-112.