

INVESTIGACIÓN CLÍNICA APLICADA

## ANÁLISIS ACÚSTICO DE LA VOZ: FIABILIDAD DE UN CONJUNTO DE PARÁMETROS MULTIDIMENSIONALES

J. GONZÁLEZ\*, T. CERVERA\*\*, J. L. MIRALLES\*\*

\*UNIVERSIDAD JAUME I DE CASTELLÓN. \*\*UNIVERSITAT DE VALÈNCIA.

### RESUMEN

La fiabilidad del análisis acústico de la voz es una cuestión relevante poco estudiada todavía. Se presenta aquí un estudio correlacional sobre la fiabilidad de los parámetros de la voz en una muestra de 148 adultos sanos. El análisis se llevó a cabo mediante el MDVP- Multi-Dimensional Voice Program instalado en una unidad de CSL-Computerized Speech Lab de Kay Elemetrics. Por cada sujeto se obtuvieron 29 parámetros de voz a partir de dos registros de la vocal /a/ sostenida. Se ofrecen los resultados generales separados por géneros y la fiabilidad test-

retest obtenida en cada par de medidas. Los datos muestran una alta estabilidad intra-sujeto en los parámetros de Frecuencia Fundamental; estabilidad aceptable en los parámetros de Perturbación de Frecuencia y de Amplitud, Ruido, Subarmónicos e Irregularidad de la voz; y una consistencia muy baja en los parámetros de Temblor. Los parámetros relacionados con el *shimmer* son más fiables que los relacionados con el *jitter*. De acuerdo con los resultados obtenidos se extraen conclusiones útiles para la práctica clínica.

**PALABRAS CLAVE:** Análisis acústico. Fiabilidad. Parámetros de voz. MDVP.

### ABSTRACT

#### ACOUSTIC VOICE ANALYSIS: RELIABILITY OF A SET OF MULTI-DIMENSIONAL PARAMETERS

A correlational study of the reliability of acoustic voice parameters was made of 148 healthy adults. Acoustic analysis was performed with MDVP- Multi-Dimensional Voice Program implemented in a CSL-Computerized Speech Lab of Kay Elemetrics. A set of 29 voice parameters were obtained from two samples of sustained vowel /a/ recorded from each subject. General results separated by sex are showed and the test-retest re-

liability in each pair of measures was calculated. Data show a high intra-subject stability of Frequency Fundamental parameters; acceptable stability in parameters of Frequency and Amplitude Perturbation, Noise, Subharmonics and Voice irregularities; and a very low consistency in Tremor parameters. Parameters related with shimmer were more reliable than parameters related with jitter. According to results several conclusions are reported.

**KEY WORDS:** Acoustic analysis. Reliability. Voice parameters. MDVP.

**Correspondencia:** J. González. Departamento de Psicología Básica, Clínica y Psicobiología. Universidad Jaume I. 12080 Castellón  
E-mail: gonzalez@psb.uji.es  
**Fecha de recepción:** 17-9-2001  
**Fecha de aceptación:** 3-1-2002

## INTRODUCCIÓN

El análisis acústico de la voz ha alcanzado un importante desarrollo en los últimos tiempos gracias, entre otras razones, al progreso y difusión experimentados por los medios informáticos que lo hacen posible. Entre sus ventajas destaca el ser un método no invasivo de evaluación de la voz y el ofrecer la oportunidad de objetivizar la evaluación en unos parámetros numéricos. Uno de los problemas principales en el diagnóstico perceptivo de la voz por el oído del clínico es que el sistema auditivo humano está preparado fundamentalmente para percibir la voz o el habla como un todo integrado, lo cual es altamente beneficioso desde el punto de vista de la comunicación lingüística. Ahora bien, esta capacidad se ve limitada cuando se trata de tomar conciencia de componentes acústicos individualizados que, sin embargo, son relevantes desde una perspectiva clínica<sup>1</sup>. En muchas ocasiones existe dificultad en determinar por un procedimiento exclusivamente perceptivo el origen de ciertas anomalías o particularidades de la voz. Por ejemplo<sup>1</sup>, algunos rasgos del tono o *pitch* son más el producto de las resonancias del tracto vocal que de la frecuencia de vibración de las cuerdas vocales. La hipernasalidad percibida en una voz puede ser la consecuencia de una desincronización en los tiempos de oclusión velar antes que una oclusión incompleta. Es decir, un mismo atributo o alteración de la calidad vocal puede tener su origen en subsistemas distintos difícilmente aislables por la mera audición. En otras ocasiones, una adecuada percepción no puede matizarse con el grado de precisión que ofrece una medida numérica. Así, en una voz percibida como soplada puede establecerse el grado de aspiración, o "breathiness", a través del parámetro pertinente (índice de turbulencia de voz). En este sentido, junto a la evaluación subjetiva por parte del clínico experimentado, el diagnóstico se enriquece y gana precisión cuando se complementa con la medida objetiva de parámetros relevantes de la voz. Las ventajas de ello se traducen en una mayor objetividad en el informe o la comunicación de los datos y en una mayor exactitud en la evaluación del progreso terapéutico, especialmente cuando éste es lento.

Para poder evaluar la voz disfónica a través de parámetros numéricos, es importante definir previamente la voz normal o no disfónica y disponer de valores normativos de comparación. Como algunos autores señalan<sup>2</sup>, en nuestro ámbito geográfico son muy escasos los estudios llevados a cabo en este sentido. Por otra parte, algunos valores

pueden ser dependientes de los algoritmos usados por el software específico que los calcula<sup>3,4</sup>, lo que hace más necesario, si cabe, el disponer de normas específicas de los principales programas empleados en la clínica.

Junto a esta necesidad, una propiedad que debe disponer todo instrumento de medida es el de una fiabilidad adecuada que permita confiar en la estabilidad de los valores obtenidos. En la extracción de parámetros de la voz no podemos esperar una coincidencia absoluta entre dos medidas sucesivas del mismo individuo y, dada la enorme variabilidad de la voz humana, tanto entre individuos como intraindividualmente, es aceptable cierta variación siempre que ésta se mantenga dentro de ciertos límites. No obstante, la robustez y validez clínica de los parámetros descansa necesariamente, como condición previa, en el grado de consistencia de sus valores. Dos medidas repetidas de la voz del mismo individuo en las mismas condiciones de registro deberían ser lo suficientemente semejantes para que podamos confiar en ellas.

El trabajo que aquí se presenta se centra en uno de los instrumentos más completos y utilizados en el análisis acústico de la voz<sup>5-7</sup>: el MDVP o Multi-Dimensional Voice Program, de Kay Elemetrics<sup>8,9</sup>. Este sistema calcula rápidamente hasta un total de 33 medidas de la función vocal pertenecientes a distintas dimensiones y las representa gráficamente en relación a valores normativos que el programa incorpora. A través de esta representación se visualiza fácilmente la significación clínica de un amplio conjunto de parámetros, gran parte de los cuales son de uso estándar en la clínica de la voz, pero otros son novedosos y propuestos recientemente<sup>9</sup>.

Kent et al. en un trabajo<sup>7</sup> muy reciente sobre el empleo del MDVP en el estudio de la voz en 32 pacientes disártricos de diversa etiología (cerebrovascular, parkinsoniana, cerebelar y otras), concluyen con el reconocimiento de la utilidad del instrumento en el habla disártrica y encuentran que algunos parámetros parecen poseer especial valor en la descripción de la calidad vocal de este grupo de pacientes. En el mismo estudio se evalúa la fiabilidad del MDVP a través de la comparación de pares de medidas dentro de cada sujeto. Observando las discrepancias intrasujeto y poniéndolas en relación con la significación clínica de los datos, los autores consideran que la fiabilidad es relativamente buena si bien "cualquier conclusión en este aspecto debe ser hecha con la apropiada comparación de la variabilidad ensayo-a-ensayo en sujetos control neurológicamente normales".

Por otra parte, los datos normativos que acompañan al MDVP son de origen americano y corresponden a un número muy limitado de casos (siete varones y ocho mujeres)<sup>8</sup>. Para llenar esa laguna, en este trabajo se proporcionan los resultados obtenidos en una muestra amplia de voces sanas pertenecientes a adultos jóvenes de ambos géneros, al tiempo que se evalúa la fiabilidad test-retest de las medidas mediante la comparación de dos registros de voz consecutivos obtenidos en cada individuo.

## MATERIAL Y MÉTODOS

### Sujetos

Se han analizado dos muestras consecutivas de voz de un total de 148 sujetos (62 hombres y 86 mujeres) de la Comunidad Valenciana con un edad media de 23,32 años (sd: 5,33) y rango comprendido entre 20 y 43 años. La mayor parte de los sujetos eran bilingües - castellano/valenciano parlantes - y todos refirieron la ausencia de cualquier trastorno de la voz, habla o audición. Los sujetos fueron seleccionados al azar de entre los cursos de Psicología y otras licenciaturas. El porcentaje de fumadores era del 48,6% (25,6% fumaban 1-10 cigarrillos diarios; 23%, más de 10 cigarrillos diarios) (ver NOTA 1)

### Procedimiento

La grabación se llevó a cabo en una sala acústicamente aislada de los laboratorios de la Universidad Jaume I de Castellón, a través de un micrófono cardiode SHURE SM58 de alta calidad dispuesto a una distancia aproximada de 10-12 cm de los labios y un ángulo de 30-45°. Las muestras consistieron en dos registros consecutivos, espaciados entre sí por un minuto de descanso, de la fonación sostenida de la vocal /a/ durante al menos tres segundos en un tono e intensidad confortables para el sujeto. La señal de voz era directamente introducida y digitalizada en el equipo de análisis CSL (*Computerized Speech Lab*) de Kay Elemetrics<sup>10</sup>, donde se encontraba instalado el programa MDVP Multi-Dimensional Voice Program modelo 4305B<sup>8</sup>. Este programa trabaja únicamente con dos frecuencias de muestreo: 25 y 50 kHz. La digitalización se realizó con una frecuencia de muestreo de 50 kHz, que es la recomendada para las vocales sostenidas, y una cuantización de 16 bits. En cada muestra se seleccionó para el análisis

un fragmento de 3 segundos de la parte más estable de la señal, desechándose en todo caso los primeros 250 milisegundos, en los que el ataque de la voz confiere inestabilidad a la señal.

### Análisis

El análisis acústico se realizó de forma automatizada por medio del programa MDVP, obteniéndose por cada muestra un total de 29 parámetros clasificados en las siguientes dimensiones: Parámetros de Frecuencia Fundamental, de Perturbación de Frecuencia, de Perturbación de Amplitud, Parámetros de Ruido, Parámetros de Temblor, de Componentes Subarmónicos, Parámetros de Aperiodicidad y de Interrupción de la voz (ver Apéndice).

## RESULTADOS Y DISCUSIÓN

En la Tabla 1 se presentan las medias, desviaciones típicas y rango de valores de los parámetros del *Multi-Dimensional Voice Program* (MDVP) de Kay obtenidos en la muestra de 148 sujetos. En la Tabla 2 se ofrecen separados por géneros, indicándose aquellos en los que surgen diferencias significativas entre varones y mujeres. No se incluyen los resultados de los parámetros de Interrupción de la voz ya que, al ser voces normales, todos los sujetos obtienen NVB y DVB con valor cero en las dos muestras, tratándose de fonaciones sostenidas sin interrupción vocal.

### Frecuencia fundamental

La media de la frecuencia fundamental ( $F_0$ ) es de 120 Hz para los varones y 200 Hz para las mujeres, con importantes variaciones en ambos grupos. Como cabía esperar, aparecen diferencias entre géneros para todos los parámetros relacionados con la frecuencia fundamental ( $F_{hi}$ ,  $F_{lo}$ , STD y PFR), siendo en todo caso la variabilidad mayor en el grupo de las mujeres, como así lo indican los valores de STD y PFR y las desviaciones típicas de  $F_0$ ,  $F_{hi}$ , y  $F_{lo}$ . El rango de valores que presenta la frecuencia fundamental ( $F_0$ ) se extiende entre 83 a 153 Hz para los varones y 158 a 274 Hz para las mujeres.

### Perturbación de la frecuencia

Dentro de los parámetros de perturbación de frecuencia, el *Jitter*, o el promedio de variación entre períodos, medido en microsegundos (millo-

**Tabla 1: Resultados de los Parámetros del MDVP (*Multi-Dimensional Voice Program*) obtenidos al analizar la fonación sostenida de la vocal /a/ en una muestra de 148 sujetos de ambos géneros**

Parámetros MDVP		Total (N = 148)		
		Media	ds	rango
Frecuencia Fundamental	Fo (Hz)	166	44,49	83 - 274
	Fhi (Hz)	176	49,57	85 - 303
	Flo (Hz)	156	41,85	70 - 263
	STD (Hz)	2,50	5,55	0,73 - 68,21
	PFR (semit.)	3,03	2,13	1 - 21
Perturbación de Frecuencia	Jita (µs)	52,35	33,32	5,84 - 192,23
	Jitt (%)	0,83	0,53	0,15 - 3,83
	RAP (%)	0,49	0,33	0,07 - 2,28
	PPQ (%)	0,49	0,31	0,08 - 2,26
	sPPQ (%)	0,72	0,36	0,30 - 3,61
	vFo (%)	1,48	3,17	0,60 - 39,26
Perturbación de Amplitud	ShdB (dB)	0,34	0,14	0,11 - 0,91
	Shim (%)	3,86	1,59	1,33 - 9,58
	APQ (%)	2,95	1,12	1,09 - 7,06
	sAPQ (%)	5,02	1,78	1,87 - 11,06
	vAm (%)	14,39	6,82	4,75 - 47,83
Ruido	NHR	0,13	0,02	0,06 - 0,24
	VTI	0,05	0,01	0,02 - 0,10
	SPI	8,54	4,43	2,23 - 29,71
Temblor	Fftr (Hz)	2,87	2,71	1,01 - 18,18
	Fatr (Hz)	2,72	1,55	1,05 - 7,54
	FTRI (%)	0,46	0,82	0,07 - 10,06
	ATRI (%)	4,60	2,58	0,89 - 13,85
Subarmónicos	DSH (%)	0,61	2,82	0 - 24,74
	NSH	0,57	2,71	0 - 24
	NSH	0,57	2,71	0 - 24
Aperiodicidad	DUV (%)	0,28	2,07	0 - 24,49
	NUV	0,28	2,03	0 - 24

nésimas de segundo) ciclo a ciclo, ha sido uno de los más empleados en la clínica y cuenta con una larga tradición en la literatura científica. El valor medio de todos los sujetos se sitúa en 52,35 microsegundos (ms) con un amplio rango comprendido entre 5,84 y 192,23 ms, si bien el 75% de la muestra se coloca por debajo de 66 ms y el 90% por debajo de 95 ms. Estos valores son ligeramente inferiores a los ofrecidos en el manual del MDVP como normativos (83,2 ms) y más en consonancia con los que aporta la literatura anterior. Por otra parte, como el *jitter* guarda relación con la frecuencia fundamental de la voz, es más inte-

resante considerarlo de forma relativa a la misma, en términos porcentuales (*Jitt*). La muestra total de sujetos presenta un promedio de 0,83% de variabilidad respecto al período fundamental, siendo ésta significativamente mayor en las mujeres (0,94%) que en los hombres (0,68%). Estos datos se sitúan entre los obtenidos en otras investigaciones. Hollien et al.<sup>11</sup> obtienen 0,48% para varones con una frecuencia fundamental Fo en torno a 102 Hz, y 0,76% para varones con Fo en torno a 142 Hz. Orlikoff<sup>12</sup> obtiene 0,42% en varones de 26-33 años.

Es creciente la tendencia a usar medidas de

**Tabla 2: Resultados de los Parámetros del MDVP (*Multi-Dimensional Voice Program*) separados por géneros**

Parámetros MDVP		Varones (N = 62)			Mujeres (N = 86)		
		Media	ds	rango	Media	ds	rango
Frecuencia Fundamental	Fo***	120	14,70	83 - 153	200	24,38	158 - 274
	Fhi***	125	16,03	85 - 167	213	27,90	166 - 303
	Flo***	115	14,20	81 - 148	186	27,43	70 - 263
	STD***	1,35	0,54	0,73 - 4,46	3,34	7,17	1,16 - 68,21
	PFR**	2,48	1,08	1 - 7	3,42	2,57	2 - 21
Perturbación de Frecuencia	Jita*	57,53	33,43	20,57 - 167,39	48,61	32,92	5,84 - 192,23
	Jitt**	0,68	0,39	0,25 - 2,14	0,94	0,59	0,15 - 3,83
	RAP***	0,39	0,24	0,14 - 1,31	0,57	0,37	0,07 - 2,28
	PPQ**	0,40	0,22	0,15 - 1,21	0,55	0,35	0,08 - 2,26
	sPPQ	0,68	0,23	0,38 - 1,52	0,75	0,42	0,30 - 3,61
	vFo	1,12	0,41	0,60 - 3,26	1,73	4,14	0,64 - 39,26
Perturbación de Amplitud	ShdB	0,33	0,14	0,11 - 0,74	0,34	0,15	0,11 - 0,910
	Shim	3,82	1,58	1,33 - 8,33	3,89	1,61	1,33 - 9,58
	APQ	3,06	1,17	1,34 - 7,06	2,87	1,08	1,09 - 5,97
	sAPQ	4,87	1,66	2,69 - 9,31	5,13	1,86	1,87 - 11,06
	vAm**	12,38	6,03	4,75 - 31,43	15,83	7,02	5,89 - 47,83
Ruido	NHR**	0,14	0,02	0,10 - 0,23	0,13	0,02	0,06 - 0,24
	VTI	0,05	0,02	0,02 - 0,10	0,05	0,01	0,02 - 0,09
	SPI	8,78	3,28	2,87 - 17,51	8,36	5,12	2,23 - 29,71
Temblor	Fftr	2,71	2,61	1,01 - 12,50	2,98	2,79	1,01 - 18,18
	Fatr	2,72	1,83	1,05 - 7,54	2,72	1,34	1,07 - 6,78
	FTRI	0,39	0,18	0,13 - 0,86	0,51	1,06	0,07 - 10,06
	ATRI*	3,98	2,50	0,92 - 13,85	5,04	2,56	0,89 - 11,17
Subarmónicos	DSH*	0,07	0,31	0,00 - 2,00	1,01	3,65	0,00 - 24,74
	NSH*	0,06	0,31	0 - 2	0,94	3,51	0 - 24
Aperiodicidad	DUV	0,55	3,14	0,00 - 24,49	0,10	0,52	0,00 - 3,16
	NUV	0,53	3,07	0 - 24	0,09	0,50	0 - 3

(\*) Diferencia significativa en la prueba Mann-Whitney con  $p < 0,05$ . (\*\*) Ídem  $p < 0,01$ . (\*\*\*) Ídem  $p < 0,001$ .

perturbación de frecuencia con un cierto grado de suavización, que no sean tan sensibles a posibles alteraciones erráticas entre dos ciclos consecutivos, o a los errores que pueda cometer el algoritmo de extracción de periodos. Las preferencias se decantan por el promediado de 3 periodos consecutivos (RAP), o 5 periodos (PPQ). Cuando se aplican estos suavizados los valores se rebajan sensiblemente, ofreciendo 0,49 para ambos parámetros en nuestra muestra total, inferiores a los datos normativos del manual del MDVP (0,68 para RAP y 0,84 para PPQ). Nuestro dato es superior

al que encuentran en España con otro programa (Dr Speech Science) Fernández Liesa et al<sup>2</sup> en una muestra de 154 adultos no fumadores (PPQ=0,23), e incluso al del grupo de fumadores del estudio de Damborenea et al.<sup>13</sup> (PPQ=0,27). También es superior al del estudio de Walton y Orlikoff<sup>14</sup> (RAP=0,28) y está más próximo a los valores encontrados por Preciado et al.<sup>15, 16</sup> en su grupo control de 64 docentes españoles (RAP=0,35), o los trabajos de Dwire y McCauley<sup>17</sup> en sujetos norteamericanos (RAP=0,38 para varones y RAP=0,89 para mujeres, ambos de 18-25 años de

edad) o Takahashi y Koike<sup>18</sup> para japoneses varones (RAP =0,57) y mujeres (RAP=0,61). Las diferencias significativas entre géneros encontradas en nuestros parámetros de perturbación frecuencial *Jita* (test Mann-Whitney,  $p<0,05$ ), *Jitt* ( $p<0,01$ ), RAP ( $p<0,001$ ), y PPQ ( $p<0,01$ ), apunta en la línea de lo declarado por Baken y Orlikoff en su reciente revisión de 2000<sup>1</sup>: "la evidencia preliminar señala la posibilidad de que las mujeres adultas podrían tener más *jitter* vocal que los hombres, al menos para algunas vocales. Afortunadamente, las diferencias observadas entre los géneros no son lo suficientemente grandes como para dar lugar a errores diagnósticos".

Si el promediado se realiza a través de 55 períodos consecutivos (sPPQ), la variabilidad frecuencial no alcanza significación entre los géneros. De nuevo, el promedio (0,72%) es inferior al umbral ofrecido por el manual (1,02%). Tampoco surgen diferencias inter-género en la variación de largo recorrido de Fo, medida a través del parámetro vFo.

### Perturbación de la amplitud

Los parámetros que miden la perturbación de amplitud a través de los ciclos de la voz (ShdB, Shim, APQ, sAPQ, y vAm) tienen, al igual que sus homólogos frecuenciales, una relevante significación clínica. Sus valores se presentan en las Tablas 1 y 2, sin diferencias estables entre géneros, a excepción de vAm. Los valores medios de *shimmer* absoluto y relativo a través de los 148 sujetos (0,34 dB y 3,86%, respectivamente) son muy semejantes a los valores normativos del manual (0,31 dB y 3,81%). En nuestra muestra el 75% de los sujetos tienen ShdB y Shim inferiores a 0,39 dB y 4,54% respectivamente; y el 90% está por debajo de 0,53 dB y 5,98%. El suavizado matemático a través del promedio de 11 períodos consecutivos, o de 55 períodos, arrojan valores medios de APQ=2,95% y sAPQ=5,02%, respectivamente. La variación de la amplitud a largo plazo en la fonación sostenida de la vocal se mide con el parámetro vAm, como el cociente, en términos porcentuales, entre la desviación típica y el promedio de la variación de amplitud ciclo-a-ciclo. En este parámetro, la diferencia entre vAm=12,38% para los varones y vAm=15,83 % para las mujeres es significativa (test Mann-Whitney,  $p<0,001$ ), presentando esta última una mayor variación de la amplitud a lo largo de la fonación sostenida. A efectos comparativos con sujetos españoles, conviene recordar que los sujetos controles de Preciado et al.<sup>15</sup> dan un valor medio de APQ =5,01%; mientras que la

muestra de no fumadores de los trabajos de Fernández<sup>2</sup> y Damborenea<sup>13</sup> arrojan APQ=2,02%, y los fumadores de este último estudio, 2,56%. En trabajos extranjeros, el "clásico" de Horii<sup>19</sup> con 31 hombres y 20 mujeres obtiene *shimmers* de 0,47 dB y 0,33 dB respectivamente para la vocal /a/, sin diferencias significativas entre los géneros. Orlikoff y Kahane<sup>20</sup> demuestran que el *shimmer* se reduce cuando la intensidad de la voz aumenta; sus datos promedio para voces masculinas a una intensidad moderada, equivalente a nuestro nivel confortable, se sitúan en 0,31 dB.

### Parámetros de ruido

Los parámetros basados en la relación entre la energía armónica y la energía de ruido tienen una extensa aplicación en la clínica de la voz, por su estrecha relación con muchas disfonías<sup>21-23</sup>. Por otra parte, son muy sensibles a los sistemas de registros empleados ya que éstos pueden introducir niveles de ruido ajenos a los que comportan la propia voz, alterando los resultados. La formulación concreta de estos parámetros varía de unos algoritmos a otros: HNR o *Harmonic-to-noise ratio*, o su recíproco, NHR *Noise-to-Harmonic ratio*, como es el caso del programa MDVP. El promedio de NHR a través de los 148 sujetos es 0,13 (es decir, la energía del ruido equivale a un 13% de la energía armónica); el 70% de los sujetos tienen valores igual o inferiores a 0,14; el 90% de los sujetos se sitúa por debajo de 0,16. Estos datos son ligeramente inferiores al 0,19 dado como umbral de normalidad por el manual del MDVP. La exigua diferencia que surge en nuestra muestra entre varones (0,14) y mujeres (0,13), es significativa (test Mann-Whitney,  $p<0,01$ ) porque aparece de forma bastante sistemática a través de los individuos, si bien no tiene significación diagnóstica.

NHR incluye globalmente toda la energía inarmónica, o de ruido, que se produce en la fonación desde orígenes diversos: irregularidades en los períodos vibratorios, subarmónicos, ruido por turbulencia, etc. El parámetro VTI-Voice *Turbulence Index*, o Índice de Turbulencia de Voz, pretende medir selectivamente el ruido que se produce por la turbulencias del aire cuando el cierre de las cuerdas vocales es incompleto o laxo. Como la mayor parte de los componentes de este ruido son de alta frecuencia, el parámetro sólo considera la energía de ruido en la región 2800-5800 Hz. Los datos obtenidos en nuestro estudio arrojan un promedio de 0,05; el 75% de los sujetos obtiene puntuaciones iguales o inferiores a 0,06; el 90% igual o inferior a 0,07. Estos datos no se apartan del

umbral de normalidad del MDVP, que lo sitúa en 0,06.

El SPI- *Soft Phonation Index*, o Índice de Fonación Blanda, es otro parámetro original de Kay Elemetrics, propuesto por primera vez en esta versión de 1993 del programa MDVP. Da una medida del predominio de componentes armónicos de baja frecuencia y se calcula dividiendo la energía armónica en el rango 70-1600 Hz entre la energía armónica en el rango 1600-4500 Hz. No se trata, por tanto, de un parámetro de ruido, pero el MDVP lo agrupa con los anteriores por la semejanza de su fórmula de cálculo y aquí mantenemos el mismo enfoque clasificatorio. El umbral de normalidad, 14,12, proporcionado por el manual del MDVP a partir de tan sólo quince sujetos se sitúa ligeramente por encima de nuestros datos. El promedio de los 148 sujetos es de 8,54; el 75% de los sujetos puntúa por debajo de 10,69; y el 90% por debajo de 13,53. Tanto en SPI como VTI no hay diferencias entre géneros.

#### Parámetros de temblor

Los parámetros de temblor, tanto frecuenciales como de amplitud, (Fftr, Fatr, FTRI, ATRI) dan valores numéricos en sólo 129 del total de 148 sujetos; en los restantes el MDVP arroja valores ausentes al no poder extraer un patrón cíclico de modulación de la voz. Salvo en ATRI, no hay diferencias entre varones y mujeres. El MDVP ofrece umbrales de normalidad en dos de ellos, FTRI=0,95% y ATRI= 4,37%, que no están lejos de los promedios obtenidos en nuestra muestra.

#### Subarmónicos

Otro rasgo de gran relevancia en el estudio clínico de la voz es la presencia de componentes subarmónicos de baja intensidad situados entre los armónicos<sup>24, 25</sup>. Esta alteración acústica suele aparecer en disfonías causadas por pólipos o edemas de Reinke, en las que pueden producirse irregularidades asimétricas en la oscilación de las cuerdas vocales y la aparición de dos o más patrones simultáneos de vibración que dan lugar a la bifonación o diplofonía. El MDVP proporciona dos parámetros, el NSH, o Número de componentes Subarmónicos, y el DSH o Grado de Subarmonicidad, consistente en la relación porcentual entre los componentes subarmónicos y los componentes armónicos de Fo. En principio la voz normal debería tener valores nulos en ambos parámetros, pero existe un pequeño porcentaje de personas, principalmente mujeres, cuya voz, sin ser patológica,

presenta componentes subarmónicos. De los 148 sujetos de nuestra muestra, 130 no tienen ningún subarmónico en su voz, mientras que 18 sujetos - 15 mujeres y 3 varones- presentan un número variable de los mismos. Esto da lugar a los valores promedios de NSH=0,57 y DSH=0,61%. Es evidente que en un parámetro de esta naturaleza el valor medio no es una adecuada aproximación desde el punto de vista clínico, pero se incluye en la tabla general con el fin de estudiar su fiabilidad test-retest.

#### Interrupción de la periodicidad

Las mismas consideraciones deben hacerse en relación con los índices NUV y DUV. Se refieren respectivamente al número y porcentaje de segmentos sordos en los que la señal deja de ser periódica y el algoritmo de cálculo no puede extraer la frecuencia fundamental. En voces normales, y al tratarse de la fonación sostenida de una vocal, no deberían arrojar valores distintos de cero. En nuestra muestra, 10 sujetos presentan alguna interrupción accidental en su fonación lo que da lugar a valores no nulos y, en consecuencia, los promedios son también distintos de cero.

#### Fiabilidad de las medidas

Más allá de la obtención de valores normativos, el objetivo principal de este trabajo ha sido evaluar el grado de fiabilidad de las medidas obtenidas, estudiando la estabilidad de los parámetros en una prueba test-retest mediante el análisis de dos muestras consecutivas de voz tomadas en cada sujeto. En la Tabla 3 se ofrecen por cada parámetro la correlación de Pearson obtenida entre ambas medidas a través de todos los sujetos. También, siguiendo la metodología de Kent, Vorperian y Duffy<sup>7</sup>, se han extraído las discrepancias, o diferencias en valores absolutos entre ambas medidas, y en la tabla se ofrecen la media, desviación típica y el rango de la distribución de esas discrepancias a través de los sujetos.

Los parámetros de la frecuencia fundamental (F0, Fhi, Flo) muestran una gran consistencia a través de las dos muestras analizadas, con correlaciones superiores a 0,90 y pequeñas discrepancias entre cada par de valores. Cuando se pide la fonación sostenida de una vocal a un tono e intensidad confortable, parece que cada sujeto tiene una frecuencia fundamental de la voz característica que varía poco de un ensayo a otro. Así el promedio de discrepancia de Fo es sólo de 3,53 Hz,

**Tabla 3: Fiabilidad test-retest de los Parámetros del MDVP (*Multi-Dimensional Voice Program*). Correlaciones de Pearson y Diferencias absolutas entre los valores obtenidos por dos muestras de voz de cada sujeto**

Fiabilidad de los Parámetros MDVP (N = 148)					
		Correlación		Diferencia Absoluta	
		1ª-2ª muestra		1ª-2ª muestra	
		r	Media	sd	rango
Frecuencia Fundamental	Fo (Hz)	0,990	3,53	5,52	0 - 52,28
	Fhi (Hz)	0,976	6,77	8,68	0 - 46,11
	Flo (Hz)	0,940	7,28	12,54	0 - 124,62
	STD (Hz)	0,216	0,95	5,35	0 - 65,24
	PFR (semit.)	0,428	1,05	1,76	0 - 15
Perturbación de Frecuencia	Jita (ms)	0,627	17,33	21,93	0 - 131,88
	Jitt (%)	0,685	0,27	0,34	0 - 2,71
	RAP (%)	0,686	0,17	0,21	0 - 1,67
	PPQ (%)	0,704	0,16	0,19	0 - 1,46
	sPPQ (%)	0,492	0,17	0,29	0 - 2,90
	vFo (%)	(NS) 0,124	0,55	3,11	0 - 37,95
Perturbación de Amplitud	ShdB (dB)	0,740	0,06	0,08	0 - 0,53
	Shim (%)	0,738	0,70	0,89	0 - 6,07
	APQ (%)	0,749	0,50	0,59	0 - 3,64
	sAPQ (%)	0,605	1,13	1,04	0 - 6,26
	vAm (%)	0,489	4,68	4,48	0 - 24,57
Ruido	NHR	0,649	0,01	0,01	0 - 0,08
	VTI	0,607	0,01	0,01	0 - 0,03
	SPI	0,848	1,66	1,81	0 - 12,61
Temblor	Fftr (Hz)	0,301	2,06	2,55	0 - 13,44
	Fatr (Hz)	(NS) 0,067	1,93	1,94	0 - 12,26
	FTRI (%)	(NS) 0,018	0,20	0,80	0 - 9,82
	ATRI (%)	(NS) 0,201	2,58	2,43	0 - 11,79
Subarmónicos	DSH (%)	0,780	0,50	1,74	0 - 13,00
	NSH	0,781	0,47	1,68	0 - 13
Aperiodicidad	DUV (%)	0,800	0,28	1,60	0 - 18,42
	NUV	0,825	0,26	1,56	0 - 18

(NS): correlación no significativa.

lo que supone un 2,12% del valor medio de Fo en todos los sujetos (166 Hz). Por otra parte, la distribución de las discrepancias es muy asimétrica, siendo relativamente unos pocos los que ofrecen dos fonaciones con tonos muy dispares, al tiempo que la mayoría muestran una alta estabilidad tonal: de los 148 sujetos, 109 presentan una discrepancia inferior a 4Hz entre la primera y la segunda

grabación. La estabilidad, sin embargo, cae notablemente cuando se considera la desviación típica (STD) de la frecuencia de cada muestra, o sus rangos fonatorios (PFR).

Entre los parámetros de Perturbación de Frecuencia, los valores de *jitter* (*Jita*, *Jitt*), RAP y PPQ muestran una estabilidad moderada con correlaciones comprendidas entre 0,62 y 0,70; siendo el

Cociente de Perturbación de Tono (PPQ) el parámetro más estable. Parece que la incorporación de un factor de suavizado de 5 períodos en PPQ, propuesto por Koike et al.<sup>26</sup>, consigue dotar a la medida de una mayor estabilidad, mostrando una consistencia entre ensayos algo superior a las medidas no suavizadas (*Jitter* absoluto y relativo) o suavizadas con 3 períodos (RAP). Un factor de suavizado de 55 períodos, sin embargo, no tiene la consecuencia de más estabilidad; antes al contrario, la correlación entre los dos medidas de sPPQ cae a 0,492.

Si se observan los promedios de las diferencias absolutas entre los dos ensayos, podemos tener una idea de las consecuencias diagnósticas de la variación entre medidas. Una diferencia promedio de 17,33 ms entre dos *jitters* consecutivos es un 30% de la media general (*Jita*=57,53  $\mu$ s). En algunos sujetos con valores próximos al umbral de normalidad, puede suponer que una muestra de voz arroje un dato por debajo de ese umbral y en una segunda muestra, tomada en las mismas condiciones, dicho umbral sea superado. En términos generales, esta diferencia de 17,33  $\mu$ s no parece dramática, pero su distribución es altamente asimétrica. De los 148 sujetos, la mitad de ellos presentó diferencias inferiores a 10  $\mu$ s entre los dos medidas de *Jita*; pero en el otro extremo de la distribución, 16 sujetos obtuvieron discrepancias por encima de 40  $\mu$ s, el cual es ciertamente un valor importante. Se constata también la tendencia, lógica por otra parte, a que el *jitter* varíe más en los sujetos con valores más altos. En los otros parámetros las diferencias son de un tenor semejante: 0,27 de diferencia entre las dos medidas de *Jitt* representa el 32,5% de la media general (0,83); 0,17 es el 34,7% de la media de RAP (0,49); 0,16 es el 32,7% de la media de PPQ (0,49); 0,17 es el 23,6% de la media de sPPQ (0,72). Es decir, podemos concluir que en líneas generales dos medidas consecutivas de la Perturbación de Frecuencia van a presentar una oscilación en torno a la tercera parte de las mismas. Si además se tiene en cuenta que en ciertos sujetos esta variación va a ser bastante mayor, el corolario parece obvio: sería aconsejable que tanto en la evaluación diagnóstica de la voz como en los estudios del progreso post-tratamiento, se obtuvieran al menos dos medidas por cada paciente y se utilizará el promedio de ambas.

Por otra parte, llama la atención la escasa consistencia de los datos de Variación de la Frecuencia Fundamental medida a través de *vFo*. Dicho parámetro refleja la variación de *Fo* a largo plazo dentro de la muestra de voz analizada. A la luz de

la prácticamente nula correlación entre los dos ensayos ( $r=0,124$ , no significativa), parece que dicho comportamiento de la frecuencia fundamental es más bien de carácter episódico y dependiente de cada fonación sostenida, antes que una característica estable de la fonación de un sujeto.

Los Parámetros de Perturbación de Amplitud exhiben, en términos generales, una consistencia superior a la de los Parámetros de Perturbación de Frecuencia. Salvo en sAPQ y *vAm*, las correlaciones entre la primera y segunda muestra de voz son superiores a 0,70 y la oscilación entre ambas viene a estar por debajo de una quinta parte de la medida. Así, una discrepancia media de 0,06 dB en ShdB representa el 17,6% del valor medio para todos los sujetos (0,34); la discrepancia de 0,70 para *Shim* es el 18,1% del total; la discrepancia de 0,50 para APQ representa el 16,9% del total. De nuevo, al igual que observábamos en los parámetros de perturbación frecuencial, podemos extraer las siguientes conclusiones: a) un factor de suavizado moderado confiere a la medida mayor estabilidad, por lo que APQ es el parámetro con más fiabilidad test-retest; b) un factor de suavizado grande (55 períodos para sAPQ) no otorga más estabilidad a la medida; c) la variación de amplitud a largo plazo dentro de la fonación, o *vAm*, es el parámetro menos fiable, como así sucede con su homólogo frecuencial.

La fiabilidad de los parámetros de ruido nos obliga a diferenciar entre NHR y VTI, por una parte, y SPI, por otra. Los dos primeros presentan correlaciones inter-muestras moderadas, de 0,649 y 0,607, respectivamente, con unas discrepancias medias relativamente inferiores en NHR (un 10% del total) que en VTI (un 20% del total). Las distribuciones de estas discrepancias son bastante asimétricas. Dada la importancia diagnóstica de NHR, conviene detenerse en la distribución de sus discrepancias entre ambas muestras de voz: de los 148 sujetos, 127 obtienen discrepancias inferiores a 0,02, lo cual tiene poca significación clínica si recordamos que el promedio total de NHR es 0,13 y el umbral de normalidad propuesto por el MDVP es 0,19. De los 21 sujetos restantes, 20 varían por debajo de 0,06, y los dos con mayores discrepancias alcanzan 0,07 y 0,08, respectivamente. Es decir, aunque la correlación entre muestras de NHR es moderada, en general la disparidad entre los valores es pequeña y tiene poca significación clínica. Consideraciones similares caben efectuar respecto a VTI, o Índice de Turbulencia de Voz, parámetro nuevo propuesto por la casa Kay en un intento de objetivizar el grado de soplado (*breathiness*) de la voz producido por un cierre laxo e in-

completo de las cuerdas vocales. Por su novedad, no disponemos aún de valores normativos en la literatura profesional.

Por su parte, el SPI, o Índice de Fonación Blanda, parece un parámetro interesante y prometedor si bien, al ser muy reciente, no cuenta hasta la fecha con estudios sobre su potencialidad diagnóstica. Más allá del ámbito clínico, promete ser un indicador objetivo de cierta cualidad de timbre vocal: la fonación "blanda", o voz con predominio de los componentes armónicos graves. Pruebas perceptivas informales revelan una estrecha correspondencia entre sus resultados numéricos y la impresión cualitativa de voz "blanda" o poco estridente, lo que es un acicate para futuros estudios sistemáticos. Este hecho tiene relevancia si tenemos en cuenta la escasa relación directa que en general encontramos en la investigación sobre timbre vocal, entre medidas numéricas y la percepción cualitativa e integrada de la cualidad vocal. En este sentido, sorprende favorablemente su alta estabilidad entre muestras, con una correlación de 0,848 entre ambas, que es la máxima después de las medidas de frecuencia fundamental (Fo, Fhi y Flo).

Los parámetros de Temblor presentan una escasísima consistencia entre ensayos. A excepción de Fftr, cuya baja correlación entre muestras (0,301), resulta estadísticamente significativa, los demás parámetros obtienen correlaciones prácticamente nulas y no significativas. Tal vez en voces patológicas en las que el rasgo de temblor sea un componente importante (temblor esencial, enfermedad de Parkinson) puedan encontrarse niveles superiores de consistencia entre ensayos, pero en lo que se refiere a las voces normales los resultados obtenidos en la primera muestra de voz son muy distintos de los que se obtienen en la segunda. Esto quiere decir que los aspectos medidos por los cuatro parámetros -frecuencia y amplitud de modulación de la voz- tienen una naturaleza marcadamente episódica ligada a cada fonación particular, con una fiabilidad escasa en cuanto a la caracterización de un rasgo estable de la voz de un individuo. Esto al menos en voces no patológicas.

Los parámetros que analizan los componentes, DSH y NSH ofrecen correlaciones test-retest próximas a 0,80. De los 148 sujetos, 122 no presentan subarmónicos en ninguna de las dos muestras registradas, 10 difieren en un subarmónico entre las dos muestras, y el resto difieren en más de un subarmónico. La consistencia de los parámetros de aperiodicidad es ligeramente superior a 0,80. Aquí, 136 sujetos no tienen segmentos aperiódicos, o sordos, en ninguna muestra, 6 difieren en un segmento,

y el resto difieren en más de un segmento aperiódico.

## CONCLUSIONES

De acuerdo con los resultados obtenidos al analizar dos muestras consecutivas de voz de 148 sujetos normales con el MDVP o *Multi-Dimensional Voice Program*, pueden extraerse las siguientes conclusiones sobre la fiabilidad de las medidas.

1. Los Parámetros de Frecuencia Fundamental (Fo, Fhi, Flo) son muy consistentes a través de las dos muestras de voz, con correlaciones test-retest muy altas. Según las características fonatorias de cada sujeto, éste tiene una frecuencia de vibración preferente que tiende a reproducir en distintas fonaciones sostenidas de la misma vocal, al menos cuando se producen a un tono e intensidad confortable. Esta consistencia, sin embargo, cae drásticamente cuando se mide la variabilidad y el rango frecuencial dentro de cada fonación (STD y PFR).

2. Los Parámetros de Perturbación de Amplitud son más consistentes que los Parámetros de Perturbación de Frecuencia, superándoles en algo más de 0,10 puntos en la correlación test-retest.

3. Dada la gran variabilidad de la voz humana, ambos tipos de parámetros presentan discrepancias entre las dos muestras de cada sujeto. En términos generales, los Parámetros de Perturbación de Frecuencia varían en torno a una tercera parte de su valor, y los Parámetros de Perturbación de Amplitud varían en torno a una quinta parte de su valor. En los casos limítrofes con la voz patológica esta variación podría tener significación clínica, por lo que sería recomendable en la práctica diagnóstica registrar al menos dos medidas y obtener el promedio de ambas.

4. Un factor de suavizado moderado en el cálculo de de ambos tipos de parámetros incrementa la consistencia de los mismos; siendo el PPQ (*Pitch Perturbation Quotient*) o Cociente de Perturbación de Tono, con un factor de suavizado de 5 períodos, el más consistente de los Parámetros de Perturbación Frecuencial, y su homólogo el APQ (*Amplitude Perturbation Quotient*) o Cociente de Perturbación de Amplitud, con un factor de suavizado de 11 períodos, el más consistente de los Parámetros de Perturbación de Amplitud. Por el contrario, un factor fuerte de suavización hace bajar la consistencia, como es el caso de sPPQ (*Smoothed Pitch Perturbation Quotient*) o Cociente Suavizado de Perturbación de Tono, y de sAPQ (*Smoothed Amplitude Perturbation Quotient*) o Cociente de Perturbación de Amplitud Suavizado,

ambos con un factor de suavizado de 55 períodos. Ahora bien, estos dos parámetros miden aspectos distintos al de los demás por lo que, pese a su inferior consistencia, son útiles porque reflejan las perturbaciones a largo término de la fonación, tanto en frecuencia como amplitud vocal.

5. Los parámetros de variación frecuencial ( $vFo$ ) y de Amplitud ( $vAm$ ) a lo largo de la fonación sostenida presentan grandes discrepancias entre las dos muestras de voz, por lo que su consistencia es baja para un mismo individuo.

6. Los Parámetros de Ruido, NHR (*Noise to Harmonic Ratio*) o Razón Ruido/Armónicos y VTI (*Voice Turbulence Index*) o Índice de Turbulencia de Voz, exhiben una consistencia moderada, con correlaciones test-retest entre 0,60 y 0,70. Dada su importancia en la práctica clínica, especialmente del primero, sería aconsejable la obtención de dos medidas promediadas, en la misma línea de lo señalado anteriormente.

7. El parámetro SPI, *Soft Phonation Index*, o Índice de Fonación Blanda es un parámetro reciente que apenas cuenta con datos sobre su potencialidad diagnóstica. Sin embargo plantea interesantes posibilidades tanto en la práctica clínica como en la investigación sobre el timbre vocal, no sólo por su aparente relevancia perceptiva sino por su gran consistencia entre muestras de voz de un mismo sujeto. De hecho, se ha revelado como el parámetro más fiable después de los parámetros de frecuencia fundamental.

8. Los parámetros de Temblor ofrecen una consistencia muy baja entre las dos muestras de un mismo sujeto con voz normal. Queda abierta la posibilidad de una mayor consistencia en voces patológicas en los que el temblor constituya un componente importante de la disfonía.

9. Los parámetros de Subarmónicos y Aperiodicidad son estables en líneas generales, si bien un pequeño porcentaje de sujetos arrojan discrepancias importantes entre sus dos muestras de voz.

## AGRADECIMIENTOS

Este trabajo han sido financiado por el Proyecto P1A99-01 de la Fundació Caixa-Castelló y la Universidad Jaume I de Castellón.

## APÉNDICE

### Parámetros de Frecuencia Fundamental:

Frecuencia Fundamental Media (*Fo-Average Fundamental Frequency*) /Hz/. El promedio de la

frecuencia fundamental para todos los períodos extraídos de la onda acústica.

Frecuencia Fundamental más Alta (*Fhi-Highest Fundamental Frequency*) /Hz/. La frecuencia fundamental más alta alcanzada en todos los períodos extraídos de la onda acústica.

Frecuencia Fundamental más Baja (*Flo-Lowest Fundamental Frequency*) /Hz/. La frecuencia fundamental más baja alcanzada en todos los períodos extraídos de la onda acústica.

Desviación típica de la Frecuencia Fundamental (*STD-Standard Deviation of Fo*) /Hz/. La desviación típica de  $Fo$  dentro del rango analizado.

Rango Fonatorio en semitonos (*PFR-Phonatory Fo-Range*). Rango entre  $Fhi$  y  $Flo$  expresado en semitonos.

### Parámetros de Perturbación de Frecuencia:

Jitter Absoluto (*Jita-Absolute Jitter*) / $\mu$ s/: Da una evaluación en microsegundos (ms) de la variabilidad período-a-período de la frecuencia fundamental dentro de la muestra analizada. Este parámetro y sus derivados son ampliamente usados en la clínica e investigación sobre la voz<sup>27</sup>, siendo muy sensible a las variaciones que ocurren entre períodos consecutivos.

Jitter Porcentual (*Jitt-Jitter Percent*) /%/: Variabilidad relativa período-a-período expresado en porcentaje.

Perturbación Relativa Promedio (*RAP-Relative Average Perturbation*) /%/: Introducida por Koike et al.<sup>26</sup>, este parámetro mide la variabilidad período-a-período con un factor de suavizado de 3 períodos.

Cociente de Perturbación de Tono (*PPQ-Pitch Perturbation Quotient*) /%/: Introducido por Koike et al.<sup>26</sup>, da la variabilidad período-a-período con un factor de suavizado de 5 períodos.

Cociente Suavizado de Perturbación de Tono (*sPPQ-Smoothed Pitch Perturbation Quotient*) /%/: Da la variabilidad período-a-período con un factor de suavizado de 55 períodos.

Variación de la Frecuencia Fundamental (*vFo-Fundamental Frequency Variation*) /%/: La desviación típica relativa de la frecuencia fundamental. Refleja la variación de  $Fo$  a largo plazo dentro de la muestra analizada.

### Parámetros de Perturbación de Amplitud:

Shimmer Absoluto (*ShdB-Shimmer*) /dB/: Medida en dB de la variabilidad entre períodos consecutivos de la amplitud pico-a-pico de la onda de la muestra de voz analizada. Como en otros parámetros, las áreas de ruptura de voz, sin naturaleza periódica, son excluidas del cómputo. Este parámetro también ha sido y es ampliamente usado en la práctica clínica<sup>27</sup>.

*Shimmer* Porcentual (*Shim-Shimmer Percent*) /%/: Variabilidad relativa entre periodos consecutivos de la amplitud pico-a-pico, expresada en porcentaje.

Cociente de Perturbación de Amplitud (*APQ-Amplitude Perturbation Quotient*) /%/: Introducido por Takahashi et al.<sup>18</sup> y Koike et al.<sup>26</sup>, mide la variabilidad relativa de la amplitud pico-a-pico con un factor de suavizado de 11 periodos. Este suavizado reduce la sensibilidad a los errores producidos en la extracción de los periodos.

Cociente de Perturbación de Amplitud Suavizado (*sAPQ-Smoothed Amplitude Perturbation Quotient*) /%/: Mide la variabilidad relativa de la amplitud pico-a-pico con un factor de suavizado de 55 periodos.

Variación de Amplitud (*vAm-Peak-Amplitude Variation*) /%/: Da la desviación típica relativa, calculada período a período, de la amplitud pico-a-pico de la onda. Refleja las variaciones de amplitud de largo término a través de la muestra analizada.

#### **Parámetros de Ruido:**

Razón Ruido/Armónicos (*NHR-Noise to Harmonic Ratio*): Una evaluación general de la presencia de ruido en la señal de voz. Este ruido puede tener diversos orígenes: variaciones en la frecuencia o la amplitud, ruido por turbulencia, presencia de componentes subarmónicos, o existencia de rupturas de voz. El parámetro consiste en el cociente entre la energía inarmónica en el rango 1500-4500 Hz y la energía armónica en el rango 70-4500 Hz. Es semejante a otros parámetros que relacionan la energías de ruido y armonicidad (HNR, SNR, etc.), de gran uso en el análisis de la voz.

Índice de Turbulencia de Voz (*VTI-Voice Turbulence Index*): Cociente entre la energía inarmónica en el rango 2800-5800 Hz y la energía armónica en el rango 70-4500 Hz. Este parámetro mide el nivel relativo de energía del ruido de alta frecuencia, en un intento de calcular de modo específico el ruido de turbulencia producido por la aspiración o breathiness de la voz. Por otra parte, es un parámetro nuevo que no cuenta con datos normativos en la literatura científica.

Índice de Fonación Blanda (*SPI-Soft Phonation Index*): Cociente entre la energía armónica en el rango 70-1600 Hz y la energía armónica en el rango 1600-4500 Hz. En realidad no se trata de un parámetro de ruido, pero el MDVP lo agrupa con los anteriores por la semejanza de su fórmula de cálculo. Bastante sensible a la estructura formántica de la vocal empleada, es un indicador de fonación "blanda" en la que predominan los componentes armónicos de baja frecuencia. También es un índice de reciente aparición, por lo que apenas cuenta con datos normativos.

#### **Parámetros de Temblor:**

Frecuencia Moduladora de Temblor de Fo (*Fftr-Fo-Tremor Frequency*) /Hz/: En la oscilación periódica de la Frecuencia fundamental (Fo) de la voz a consecuencia de temblor, este parámetro muestra la frecuencia de modulación de Fo.

Frecuencia Moduladora de Temblor de Amplitud (*Fatr-Amplitude Tremor Frequency*) /Hz/: En la oscilación periódica de la Amplitud de la voz a consecuencia de temblor, este parámetro muestra la frecuencia de modulación de la misma.

Índice de Intensidad de Temblor de Fo (*FTRI-Fo-Tremor Intensity Index*) /%/: En la oscilación periódica de Fo a consecuencia de temblor, este parámetro muestra la magnitud de dicha frecuencia moduladora en respecto a la magnitud de frecuencia total de la señal analizada.

Índice de Intensidad de Temblor de Amplitud (*ATRI-Amplitude Tremor Intensity Index*) /%/: En la oscilación periódica de la Amplitud de la voz a consecuencia de temblor, este parámetro muestra el porcentaje de intensidad del temblor respecto a la amplitud total de la señal.

#### **Parámetros de Componentes Subarmónicos:**

Número de Componentes Subarmónicos (*NSH-Number of Subharmonic Segments*): Número de componentes subarmónicos encontrados durante el análisis. Se trata de un indicador de diplofonía o bifonación.

Grado de Subarmonicidad (*DSH-Degree of Subharmonics*) /%/: Relación entre los componentes subarmónicos y los componentes armónicos de Fo.

#### **Parámetros de Aperiodicidad de la Voz:**

Número de segmentos sordos (*NUV-Number of Unvoiced Segments*): Número de segmentos sin naturaleza armónica detectados durante el análisis.

Grado de sordera (*DUV-Degree of Voiceless*) /%/: Porcentaje de las áreas no armónicas en la muestra de voz.

#### **Parámetros de Interrupción de la Voz:**

Número de interrupciones de la voz (*NVB-Number of Voice Breaks*): Número de veces en que Fo resulta interrumpida en la muestra de voz analizada.

Grado de interrupción de la voz (*DVB-Degree of Voice Breaks*) /%/: Porcentaje de la longitud de las áreas que representan interrupciones de voz respecto a la longitud total de la muestra analizada.

#### **NOTA 1**

El porcentaje de fumadores de la muestra no se aparta sensiblemente de la prevalencia de la población española, pues si, según el Informe so-

bre la Salud de los Españoles: 1998 del Ministerio de Sanidad y Consumo, la prevalencia del hábito tabáquico en 1997 se situaba en el 35,7% de la

población general, los datos de ese mismo año para el grupo de edad 16-24 años son el 39,7% y para 25-44 años, 52,3%.

## REFERENCIAS

- 1.- Baken R, Orlikoff, R. Clinical measurement of speech and voice. Second Edition. San Diego, CA: Singular Publishing Group, 2000.
- 2.- Fernández R, Damborenea D, Rueda P, García E, Leache J, Campos MA, Llorente E, Naya MJ. Análisis acústico de la voz normal en adultos no fumadores. *Acta Otorrinolaringol Esp* 1999; 50(2): 134-41.
- 3.- Read Ch, Buder E, Kent R. Speech Analysis Systems: An evaluation. *J Speech Hear Res* 1992; 35: 314-32.
- 4.- Parsa V, Jamieson D. A comparison of high precision Fo extraction algorithms for sustained vowels. *J Speech Hear Res* 1999; 42: 112-26.
- 5.- Buder E. Acoustic analysis of voice quality: A tabulation of algorithms 1902-1990. En R Kent y M Ball. *Voice Quality Measurements*. San Diego: Singular Publishing Group, 2000.
- 6.- Corina J, Hilgers F, Verdonck I, Koopmans F. Acoustical analysis and perceptual evaluation of tracheoesophageal prosthetic voice. *J Voice* 1998; 12: 239-48.
- 7.- Kent R, Vorperian H, Duffy J. Reliability of the Multi-Dimensional Voice Program for the analysis of voice samples of subjects with dysarthria. *Am J Speech-Lang Pathol* 1999; 8: 129-36.
- 8.- Kay Elemetrics Corp. Multi-Dimensional Voice Program. MDVP Model 4305B. New Jersey 1993.
- 9.- Deliyiski D. Acoustic model and evaluation of pathological voice production. *Kay Elemetrics*, 1993.
- 10.- Kay Elemetrics Corp. Computerized Speech Laboratory. CSL Model 4300-B. New Jersey 1994.
- 11.- Hollien H, Michel J, Doherty E. A method for analyzing vocal jitter in sustained phonation. *J Phonetics* 1973; 1: 85-91.
- 12.- Orlikoff R. Heartbeat-related fundamental frequency and amplitude variations in healthy young and elderly male voices. *J Voice* 1990; 4: 322-28.
- 13.- Damborenea J, Fernández R, Llorente E, Naya MJ, Marín C, Rueda P, Ortiz A. Efecto del consumo de tabaco en el análisis acústico de la voz. *Acta Otorrinolaringol Esp* 1999; 50(6): 448-52.
- 14.- Walton J, Orlikoff R. Speaker race identification from acoustic cues in the vocal signal. *J Speech Hear Res* 1994; 37: 738-45.
- 15.- Preciado JA, García R, Infante JC. Análisis multidimensional de la función vocal. Estudio de casos y controles. *Acta Otorrinolaringol Esp* 1998; 49(6): 467-74.
- 16.- Preciado JA, Fernández S. El análisis digital de la señal acústica en el diagnóstico de la patología vocal. Sensibilidad y especificidad de las medidas del shimmer y del jitter. *Acta Otorrinolaringol Esp* 1998; 49(6): 475-81.
- 17.- Dwire A, McCauley R. Repeated measures of vocal fundamental frequency perturbations obtained using de Visi-Pitch. *J Voice* 1995; 9: 156-62.
- 18.- Takahashi H, Koike Y. Some perceptual dimensions and acoustical correlates of pathologic voices. *Acta Otolaryngol Suppl* 1975; 338: 1-24.
- 19.- Horii Y. Fundamental frequency perturbation observed in sustained phonation. *J Speech Hear Res* 1979; 22: 5-19.
- 20.- Orlikoff R, Kahane J. Influence of mean sound pressure level on jitter and shimmer measures. *J Voice* 1991; 5: 113-19.
- 21.- Yanahigara, N. Significance of harmonic changes and noise components in hoarseness. *J Speech Hear Res* 1967; 10: 531-41.
- 22.- Yumoto E, Gould W, Baer Th. Harmonic to noise ratio as an index of the degree of hoarseness. *J Acoust Soc Am* 1982; 71: 1554-60.
- 23.- Nieto A, Cobeta I, Gamboa FJ, Vegas A. La relación armónico/ruido y el análisis espectrográfico en la patología de abuso vocal. *Acta Otorrinolaringol Esp* 1996; 47(5): 370-76.
- 24.- Omori K, Kojima H, Kakani R, Slavik DH, Blaugrun SM. Acoustic characteristics of rough voice: Subharmonics. *J Voice* 1997; 11: 40-47.
- 25.- Núñez F, Suárez C, Muñoz C, Baragaño L, Álvarez MJ, Martínez A. Estudio espectro-gráfico de la disfonía: subarmónicos. *Acta Otorrinolaringol Esp* 2000; 51(1): 52-56.
- 26.- Koike Y, Takahashi H, Calcatera, T. Acoustic measures for detecting laryngeal pathology. *Acta Otolaryngol* 1977; 84: 105-17.
- 27.- Iwata S, von Lenden H. Pitch perturbations in normal and pathological voices. *Folia Phoniatr* 1970; 22: 117-28.