



COMPARACIÓN DE VOCES DISFÓNICAS Y VOCES NO DISFÓNICAS SEGÚN CRITERIOS DE JUECES ENTRENADOS

E. MENDOZA; J. VILA
Universidad de Granada

Resumen

El objetivo de este estudio ha sido la obtención de unos criterios perceptivos fiables y válidos aplicables a la evaluación vocal de personas que presentan disfonía profesional. Mediante un análisis y evaluación de muestras de voz de acuerdo con jueces entrenados, se han encontrado diferencias estadísticamente significativas entre voces disfónicas (diagnosticadas por especialista en laringología) y voces no disfónicas en las variables vocales de intensidad, tono, disfonía, y tres tipos de cualidad de voz con amplia aceptación (velada, ronca y áspera). La intercorrelación de las variables de las tres frases analizadas es muy elevada. A su vez, la variable intensidad correlaciona negativamente con la cualidad de voz velada, y la variable tono sigue la misma tendencia con respecto a todas las variables relativas a cualidad de voz. Se han determinado tres factores en lo que presentan saturaciones elevadas las variables relacionadas con la cualidad, tono e intensidad, respectivamente.

Abstract

The aim of this study is to obtain some reliable and valid criteria to be applied to the vocal assessment of people who have professional dysphonia. By means of an analysis and assessment of voice samples according to trained judges, it has been found statistically significant differences between dysphonic (diagnosed by otolaryngology specialist) and non-dysphonic voices in the vocal variables of intensity, pitch, dysphonia and three types of voice quality with wide acceptance (breathy, hoarse, and harsh). The intercorrelation of the variables of the three sentences used in this work is very strong. Besides, the variable intensity is negatively correlated with the quality of breathy voice, and the variable pitch has the same trend with respect to all variables connected to voice quality. Finally, it has been determined three factors for which the variables related to quality, pitch and intensity, respectively, present high saturations.

Introducción

El conocimiento de los parámetros vocales alterados en voces disfónicas es excepcionalmente complejo bajo puntos de vista, tanto metodológico como conceptual. En el campo de la evaluación vocal es tarea de gran importancia el establecimiento de unos criterios sólidos y objetivos que permitan identificar los parámetros vocales que se modifican cuando la voz adquiere carácter de disfonía.

Actualmente, determinados tipos de desórdenes vocales se pueden considerar como enfermedades profesionales, o enfermedades de relativamente larga duración adquiridas durante el ejercicio profesional y consecuencia del mismo. Un claro ejemplo de estas enfermedades se puede observar en las disfonías profesionales, o desórdenes vocales característicos de distintos ámbitos profesionales en los que

la voz se tiene que utilizar excesiva o abusivamente (cantantes, actores de teatro y profesionales de la enseñanza y la comunicación, entre otros). Éste es un problema creciente en nuestra sociedad, incrementándose en el campo de profesionales de la enseñanza en un 205 por 100, en los últimos años en nuestro país (Esteve, 1989).

Los problemas de voz profesionales se pueden caracterizar bajo el denominador común de la hiperfunción vocal, definida como el uso excesivo de la fuerza aductora durante la fonación, que puede producir o mantener determinados cambios patológicos. Es bien sabido que los parámetros de voz generalmente aceptados son los de intensidad, tono, duración y cualidad de la voz con todas sus ambigüedades. De todos ellos, tono, intensidad y duración tienen claros correlatos físicos y pueden ser cuantificados (Michel y Wendhal, 1971). Sin embar-

go, la cualidad es un parámetro excesivamente ambiguo y difícilmente cuantificable. Los términos utilizados para medir la cualidad de la voz no son parámetros descriptivos del funcionamiento laríngeo.

Actualmente existen distintos tipos de instrumentaciones y técnicas para determinar el tono de voz (Moncur y Brackett, 1974; Boone, 1983; Fairbanks, 1960; Hollien y Moore, 1966), siendo aún más fácil la determinación de la duración y la intensidad. No obstante, la cualidad presenta serias dificultades de objetivación, incluso en los perfiles más completos y actuales (Wilson y Broda-Masiuk, 1987). Es probable que el estudio de la cualidad de la voz se tenga que realizar de forma más analítica, a través de la investigación de distintos tipos de cualidad de voz. La investigación en este punto es bastante parca y poco concluyente.

Probablemente, la polémica en torno a técnicas de intervención vocal y a la eficacia relativa de cada una de ellas obedezca a la falta de clarificación de estos parámetros alterados, si no en la disfonía de forma global y abstracta, si en determinados tipos de voces disfónicas. Actualmente sabemos con mayor precisión por qué se altera la voz que en qué se manifiesta dicha alteración.

El objetivo que se pretende alcanzar con este estudio es la determinación de los parámetros vocales que diferencian las voces disfónicas de las no disfónicas, según la evaluación de jueces entrenados al respecto. Actualmente carecemos de investigaciones metodológicamente rigurosas sobre los parámetros alterados en voces disfónicas y, lógicamente, sobre la efectividad de la terapia vocal (Reed, 1980). La investigación básica en la conducta vocal es muy abundante, mientras la investigación aplicada es prácticamente inexistente.

De acuerdo con Wedin y col. (1978, pág. 103), «La cualidad de la voz es tan bien percibida por el oído humano que probablemente nunca pueda ser superada por ningún instrumento electro-acústico». El reconocimiento de la conducta vocal de uno o varios hablantes (*encoders*) por unos jueces más o menos entrenados (*decoders*) ha sido una línea de investigación tradicionalmente complementaria a los análisis acústicos objetivos, fundamentalmente el reconocimiento de emociones a través de indicadores no verbales de la conducta vocal (Soskin y Kauffman, 1961; Kramer, 1963; Scherer, 1982; Scherer y col, 1972; Burns y Beier, 1973; Sogon, 1975; Zuckerman y col, 1975; Jiménez, 1985). La metodología para llevar a cabo este tipo de estudios se puede encontrar magistralmente descrita en Rosenthal (1982), metodología que hemos intentado adaptar al reconocimiento de determinados atributos vocales en distintos tipos de voces (disfónicas y no disfónicas), siempre dentro del campo de las llamadas «disfonías profesionales».

Para llevar a cabo el análisis de voz por el procedimiento de jueces seguimos el método básico descrito por Rosenthal (1982). Una serie de hablantes caracterizados por uno o más atributos (presencia o ausencia de disfonía en nuestro caso) son observa-

dos —escuchados— por varios jueces que emiten uno o más juicios sobre los hablantes en base a una conducta vocal selectivamente presentada. La propia conducta vocal de los hablantes actúa como variable dependiente, y nuestro propósito en esta investigación era que los jueces emitieran sus criterios sobre seleccionados atributos de la voz de los sujetos para determinar si estos juicios discriminan adecuadamente la presencia/ausencia de disfonía.

Las principales dificultades para alcanzar este objetivo radican en la gran variabilidad vocal inter e intra hablante. Atkinson (1976), tras asumir la frecuencia fundamental (Fo) como el parámetro acústico más determinante en la fonación, detectó más variabilidad intrahablante que entre distintos hablantes. Pero, a pesar de éste y otros muchos obstáculos, si en esta línea de investigación se obtuviesen resultados con un mínimo nivel de claridad, el apoyo a la terapéutica vocal sería incuestionable.

Método

Sujetos

Los sujetos participantes en esta primera parte del estudio fueron 28 mujeres, profesionales de la enseñanza en activo, en los niveles de preescolar, EGB y BUP, que al menos llevaban un curso escolar completo en la enseñanza. De todas ellas, 14 presentaban problemas vocales de los que usualmente se vienen asociando a las disfonías profesionales y el abuso vocal, diagnosticado por algún médico especialista en ORL. Ellas formaron el grupo experimental. Los 14 sujetos restantes, igualmente mujeres profesionales de la enseñanza, eran hablantes normales, sin historia de problemas de voz ni audición, y formaron el grupo control de nuestro experimento. La edad de los sujetos del grupo experimental oscilaba entre 22 y 47 años (media = 34,78 y DT = 7,43). Por su parte, el rango de edad del grupo control oscilaba entre los 24 y 52 años (Media = 37,21 y DT = 9,10).

Aparatos

Se utilizó un grabador SONY, modelo TCM-5000-EV, y un micro AKG D222 EB para la recogida de muestras de voz de todos los sujetos. Las grabaciones se hicieron siempre sobre cintas de cassette TDK SF60. Las muestras de voz eran grabadas en una habitación semiinsonorizada, manteniendo el micro a una distancia aproximada de 20 cm de la boca.

Selección de frases

La selección de las frases se hizo de la forma siguiente:

Se eligieron cinco frases de una relación de veinte siguiendo un triple criterio: a) que fueran fáciles de recordar y memorizar, ya que queríamos analizar muestras de habla espontánea y no de lectura, debido a que el modelo general de entonación varía en ambas situaciones (Schultz-Coulon, 1975); b) que fueran ricas en fonemas sonoros y en contraste sonoro/sordo para que reflejasen al máximo el trabajo activo de cuerdas vocales (Quilis, 1981; Golderos y col., 1980; Laver, 1980), y c) que fuesen de contenido neutro para los sujetos, ya que si al emitir las se utiliza un modelo especial de entonación en función del contenido emocional, la configuración general de los parámetros acústicos puede verse afectada (Fairbanks y Hoaglin, 1941; Davitz, 1964; Williams y Stevens, 1969; Scherer, 1979; Jiménez, 1985).

Se eligieron también cinco refranes de entre un total de 20. En la selección de refranes sólo se tuvo en cuenta el criterio a), ya que para nuestra finalidad eran una especie de control del contenido emocional, y, al ser expresiones de la llamada «sabiduría popular», difícilmente esperábamos que fueran neutros en contenido.

En total se hizo una recopilación de diez expresiones lingüísticas —cinco frases y cinco refranes— que se ordenaron al azar y se sometieron a una evaluación mediante diferencial semántico (Osgood y Tannenbaum, 1957; Fierro, 1983) en 10 mujeres del mismo colectivo profesional.

De todas las frases elegimos una que resultó ser neutra en las escalas de evaluación, potencia y actividad (Fierro, 1983): «A LOS LOMOS DEL CABALLO VA LA YEGUA», y que, excepto la /c/ inicial de «caballo», no contiene ningún fonema sordo.

De entre los refranes elegimos dos: «CUANDO EL RIO SUENA, AGUA LLEVA», que resultó estar polarizado positivamente en las tres escalas, y «AL PAN PAN, Y AL VINO VINO», que los sujetos lo evaluaron como neutro en cada una de las escalas.

Procedimiento

Los 28 sujetos pasaron por una sesión de evaluación individual en un centro de tratamiento de problemas de habla y lenguaje. Cuando llegaban al centro de trabajo se les explicaba brevemente la finalidad del estudio y de las grabaciones que se iban a realizar, y se les pedía que leyeran varias veces, en silencio, las frases seleccionadas impresas cada una en una tarjeta, para que después pudieran grabarlas sin necesidad de leerlas. Se insistió siempre en que emitieran las frases con una intensidad de voz normal, evitando toda afectación y emocionalidad en las mismas. Una vez que decían tenerlas aprendidas, se les presentaba la tarjeta con la frase impresa durante 2 segundos y, siempre en el mismo orden, dejando un intervalo de unos 5 segundos aproximadamente entre frase y frase;

«Cuando el río suena, agua lleva»
«A los lomos del caballo va la yegua»
«Al pan pan, y al vino vino».

Cuando finalizó la fase de grabación, las muestras de voz fueron evaluadas por 21 jueces en una sesión conjunta. El grupo de jueces estaba formado por 21 alumnos de quinto curso de Psicología de la Universidad de Granada —6 varones y 15 mujeres— que habían superado satisfactoriamente el programa teórico y práctico de Terapia Vocales de la asignatura Terapéutica del Lenguaje y la Motricidad. Dicho programa práctico consistió precisamente en la emisión de juicios sobre determinados atributos de voces previamente grabadas.

Una vez reunidos a los alumnos en un aula de la facultad, se les hizo entrega de unas hojas con las instrucciones del experimento. Leídas y aclaradas las dudas surgidas, se evaluaron las voces de cinco sujetos no experimentales a modo de ensayo y demostración, y comenzó la sesión experimental propiamente dicha. Las instrucciones dadas a los jueces fueron las siguientes:

«El objetivo de este estudio es someter una serie de frases a una evaluación de jueces para comparar los juicios procedentes de sujetos —jueces— con un entrenamiento determinado en reconocimiento de distintas características vocales, con los resultados de un análisis acústico objetivo de ciertos parámetros vocales.

Las frases que tendréis que evaluar son siempre las mismas y en el mismo orden:

“Cuando el río suena, agua lleva”
“A los lomos del caballo va la yegua”
“Al pan pan, y al vino vino”

Estas frases las oiréis repetidas veces dichas por distintos sujetos, siempre mujeres, presentando algunos de ellos un problema de disfonía. El orden de las grabaciones de los distintos sujetos se ha aleatorizado.

Tras la audición de “sujeto 1”, “sujeto 2”, 3, 4, etc., aparecerá un bloque con las tres frases mencionadas, que tendréis que evaluar una a una.

Si miráis vuestra hoja de respuesta veréis en la primera columna la palabra “frases”, y debajo números del 1 al 3 repetidos cada vez y separados por una línea más gruesa. Son las tres frases correspondientes a cada uno de los sujetos. Tras oír cada frase marcaréis con una cruz la casilla correspondiente a vuestra evaluación en cada una de las escalas siguientes: intensidad, tono, disfonía, velada, ronquera y aspereza.»

Cada juez tenía que evaluar, por tanto, tres frases de cada sujeto (frase 1, frase 2 y frase 3), y seis atributos en cada frase, resultando un total de 18 variables por sujeto. Los bloques de las tres frases de cada sujeto fueron aleatorizadas y se presentaron en el orden siguiente: E-C-E-C-E-E-C-C-C-E-C-C-E-C-E-E-E-E-E-E-C-C-E-C-C-E, donde E indica sujeto del grupo experimental y C, sujeto del grupo control. Durante la sesión de evaluación los jueces estaban sentados en lugares lo suficientemente separados como para evitar cualquier intento de comunicación verbal o visual. Así mismo se les pidió que no manifestaran gestos ni señales de

asombro o hilaridad para no influir en las respuestas de los compañeros.

Escalas de evaluación

Se utilizaron las seis escalas de evaluación siguientes junto con las instrucciones definitorias de cada escala y de cada atributo a evaluar:

INTENSIDAD: Escala bipolar: 1, Muy débil; 2, Débil; 3, Normal; 4, Fuerte; 5, Muy fuerte.

TONO: Escala bipolar. Cuando hablamos de tono grave y tono agudo no podemos ofrecer unos datos-guías para su determinación. Cuando decimos que una voz es grave o aguda lo que hacemos, en definitiva, es compararla con voces de otras personas de sexo y edad similar. En nuestro caso, todas las voces corresponden a mujeres adultas de edades comprendidas entre 22 y 54 años, siendo la escala la siguiente: 1, Muy grave; 2, Grave; 3, Normal; 4, Aguda; 5, Muy aguda.

DISFONÍA: La escala de evaluación de la disfonía no es bipolar. Oscila entre 1 —voz normal, no disfónica— y 5 —disfonía muy severa—. Entendemos por disfonía «Voz anormal, a juicio del oyente, que implica tanto al tono como a la intensidad, la cualidad y, en algunos casos, la duración» (Aronson, 1985). No se os pueden aportar más datos sobre el término, ya que lo decisivo en la definición es, en definitiva, el juicio del oyente. La escala correspondiente es: 1, Voz normal (no disfónica); 2, Disfonía ligera; 3, Disfonía moderada; 4, Disfonía severa; 5, Disfonía muy severa.

VELADA: La cualidad de voz velada o aspirada es difícil de precisar. En palabras de Fairbanks (1960) se define por un excesivo escape de aire producido por un insuficiente cierre glotal y un ataque vocal aspirado. La relajación de la válvula laríngea parece estar presente también en este tipo de voz (Burkowsky y Vitali, 1988). Calificativos como «voz quebrada» y «voz temblona» han caracterizado a esta cualidad de voz (Scherer, 1982). La escala, no polar, oscila entre 1 y 5: 1, Normal; 2, Ligeramente velada; 3, Moderadamente velada; 4, Bastante velada; 5, Muy velada.

RONQUERA: Aunque en el lenguaje vulgar ronquera puede entenderse como un término equivalente a disfonía, se refiere a un tipo muy concreto de disfonía, intermedio entre voz velada y voz áspera que veremos a continuación, aunque con predominio de una sobre otra (Fairbanks, 1960). Se puede entender como «voz gutural» (Scherer, 1982) y fácilmente se puede asemejar a la voz «tomada» o algo «cascada» propia de quien, sin ningún tipo de entrenamiento vocal, pasa largo rato cantando. Por alguna razón, las cuerdas pueden tener dificultad en unirse y separarse suavemente. La escala, no polar, oscila entre 1 y 5: 1, Normal; 2, Ligeramente ronca; 3, Moderadamente ronca; 4, Bastante ronca; 5, Muy ronca.

ASPEREZA: El calificativo de áspero o duro se aplica a la voz para referirse a una emisión vocal caracterizada por un tono generalmente muy grave,

con inicio abrupto (ataque duro), esfuerzo creciente y fuerza vocal agresiva (Burkowsky y Vitali, 1988). Es el tipo de voz que frecuentemente se asocia al habla del político en el escenario de un mitin y al ejecutivo duro y agresivo. La escala de evaluación es la siguiente: 1, Normal; 2, Ligeramente áspera; 3, Moderadamente áspera; 4, Bastante áspera; 5, Muy áspera.

Resultados

Los resultados obtenidos en esta evaluación por jueces fueron los siguientes:

Fiabilidad efectiva

Los índices de fiabilidad se obtuvieron a través de análisis de varianza y la aplicación de las fórmulas propuestas por Rosenthal para la obtención de la fiabilidad efectiva ($R_{[est]}$) y la fiabilidad media de cada juez ($r_{[est]}$) (Rosenthal, 1982):

$$R_{(est)} = M. C. jueces - M. C. residual / M. C. jueces \\ r_{(est)} = M. C. jueces - M. C. residual / M. C. jueces \\ + (n-1)M. C. residual, \text{ donde } M. C. \text{ es la media de cuadrados del Análisis de Varianza.}$$

Los resultados de los coeficientes de fiabilidad para cada una de las variables fueron muy elevados, oscilando los índices de fiabilidad efectiva entre 0,9735 y 0,7578, y los índices de fiabilidad media de cada juez entre 0,6367 y 0,1296.

Análisis de varianza

En la tabla 1 se pueden observar las Medias, las Desviaciones Típicas y los valores de F con sus probabilidades asociadas de las evaluaciones de los jueces en las diferentes variables vocales en los dos grupos de sujetos. Los resultados se analizaron mediante un ANOVA unifactorial de grupos independientes —grupo de Disfónicos y grupo Control— utilizando el programa 2V del paquete estadístico BMDP (Dixon, 1983). Según los resultados del análisis, excepto en la variable intensidad de la frase primera, cuya diferencia entre los dos grupos se halla próxima a la significación estadística, los valores de «F» son muy elevados, alcanzando valores de significación estadística ($p < 0,001$) en todas las variables de todas las frases, excepto en la variable INTENSIDAD (no significativa en la frase primera y $p = < 0,01$ en las frases segunda y tercera) y en la variable ASPERA de las frases segunda y tercera ($p = < 0,01$). Podemos, por tanto, afirmar que las diferencias entre los grupos de sujetos —control y disfónico— son estadísticamente significativas en todas las variables.

TABLA 1

Medias, Desviaciones típicas, valores de F y probabilidades asociadas de las evaluaciones de los jueces para cada una de las variables de ambos grupos (Control y Disfónico)

Variable	Controles		Disfónicos		F
	Media	D.T.	Media	D.T.	
ITENS.-1	3,0444	0,6312	0,26055	0,7080	3,00
TONO-1	3,3876	0,3282	2,5373	0,4650	31,24 ***
DISFON.-1	1,6259	0,3298	2,3808	0,6994	13,34 ***
VELADA-1	1,4761	0,2654	1,9932	0,3983	16,34 ***
RONCA-1	1,2347	0,1685	2,1700	0,6914	42,18 ***
ÁSPERA-1	1,2585	0,1368	1,6259	0,3659	12,38 ***
INTENS.-2	3,2483	0,6044	2,5883	0,6231	8,09 **
TONO-2	3,2790	0,3808	2,5203	0,3761	28,12 ***
DISFON.2	1,5851	0,2158	2,3129	0,6450	16,03 ***
VELADA-2	1,4252	0,2222	1,9513	0,3888	21,84 ***
RONCA-2	1,2380	0,1333	2,1122	0,6576	27,73 ***
ÁSPERA-2	1,2720	0,1634	1,5680	0,3525	8,12 **
INTENS.-3	3,1087	0,4231	2,5271	0,5291	10,32 **
TONO-3	3,2210	0,2723	2,5579	0,2773	40,74 ***
DISFON.-3	1,4217	0,2192	2,2312	0,6845	17,76 ***
VELADA-3	1,3264	0,1594	1,8026	0,3463	21,84 ***
RONCA-3	1,2481	0,1723	2,0203	0,5821	22,65 ***
ÁSPERA-3	1,2515	0,1366	1,5338	0,3110	9,66 ***

* p = < 0,05
 ** p = < 0,01
 *** p = < 0,001

Análisis correlacional

Las puntuaciones medias de los jueces en las distintas variables se sometieron a un análisis correlacional; en primer lugar, mediante la correlación de Pearson, en segundo lugar, mediante un Análisis Factorial utilizando el método de componentes principales y rotación ortogonal.

Correlaciones de Pearson

Como puede verse la tabla 2, las puntuaciones en INTENSIDAD de cada frase correlacionan alta y positivamente con las puntuaciones de intensidad de las restantes frases (r = 0,914 y r = 0,835, respectivamente). Se contemplan también correlaciones negativas de dicha variable con «voz velada» (r = -0,647 en la frase primera, r = -0,597 en la frase segunda y r = -0,577 en la frase tercera), por lo que los jueces parecen referirse a «voz velada» cuando ésta se emite a muy baja intensidad.

La variable TONO correlaciona alta y negativamente con todas las variables que se refieren a la cualidad de la voz (disfonía como descripción general, voz velada, ronca y áspera). La correlación de la variable tono en las tres frases ha resultado positiva y muy alta.

La variable DISFONIA aparece en todas las frases

como relativamente independiente de la intensidad de emisión de las mismas, siendo su correlación, tal como ya ha quedado comentado, más alta y negativa con el tono de voz (r = -0,663, r = -0,592 y r = -0,611 en las frases primera, segunda y tercera, respectivamente). Los jueces aprecian la disfonía como muy asociada con todas las variables referentes a cualidad de voz.

En las restantes variables se observa la misma tendencia: correlación alta y positiva entre ellas. La correlación más baja se registra entre las variables «velada» y «áspera» (r = 605, r = 564 y r = 715 en frases primera, segunda y tercera, respectivamente), lo cual puede indicar la relativa independencia de ambas categorías de cualidad de voz.

Análisis factorial

Los datos procedentes de la evaluación de jueces fueron sometidos a un Análisis Factorial mediante el programa 4M del EMDP, cuyos resultados aparecen en la tabla 3. En ella se recogen los tres factores obtenidos, la saturación de las variables en cada factor y la proporción de la varianza que viene explicada por cada uno de ellos.

El primer factor, que se puede definir como factor «cualidad», explica una parte muy alta de la varianza

ANÁLISIS CORRELACIONAL. CUADRO DE INTERCORRELACIONES DE CADA UNA DE LAS VARIABLES CON TODAS LAS DEMÁS, SEGÚN EVALUACIÓN DE LOS JUECES

	INT1	TON1	DIS1	VEL1	RON1	ASP1	INT2	TON2	DIS2	VEL2	RON2	ASP2	INT3	TON3	DIS3	VEL3	RON3	ASP3
INT	1,000																	
TON1	0,487	1,000																
DIS1	-0,380	-0,663	1,000															
VEL1	-0,647	-0,721	0,831	1,000														
RON1	-0,311	-0,823	0,902	0,785	1,000													
ASP1	-0,015	-0,749	0,746	0,605	0,878	1,000												
INT2	0,914	0,562	-0,345	-0,606	-0,342	-0,066	1,000											
TON2	0,453	0,962	-0,588	-0,650	-0,772	-0,707	0,564	1,000										
DIS2	-0,321	-0,654	0,951	0,773	0,877	0,727	-0,345	-0,592	1,000									
VEL2	-0,532	-0,741	0,780	0,901	0,787	0,630	-0,597	-0,668	0,823	1,000								
RON2	-0,212	-0,757	0,867	0,720	0,954	0,838	-0,291	-0,723	0,908	0,792	1,000							
ASP2	0,178	-0,591	0,677	0,477	0,757	0,899	0,113	-0,585	0,721	0,564	0,816	1,000						
INT3	0,835	0,538	-0,354	-0,569	-0,347	-0,055	0,935	0,551	-0,352	-0,522	-0,289	0,106	1,000					
TON3	0,450	0,947	-0,594	-0,651	-0,781	-0,705	0,549	0,967	-0,595	-0,671	-0,729	-0,590	0,558	1,000				
DIS3	-0,313	-0,637	0,924	0,743	0,854	0,684	-0,328	-0,588	0,964	0,752	0,881	0,676	-0,375	-0,611	1,000			
VEL3	-0,525	-0,699	0,889	0,887	0,824	0,600	-0,544	-0,634	0,913	0,885	0,824	0,539	-0,577	-0,648	0,905	1,000		
RON3	-0,227	-0,733	0,853	0,704	0,925	0,784	-0,283	-0,706	0,899	0,747	0,967	0,765	-0,330	-0,734	0,918	0,847	1,000	
ASP3	0,089	-0,669	0,726	0,543	0,826	0,936	0,040	-0,650	0,753	0,623	0,851	0,966	0,049	-0,666	0,715	0,600	0,806	1,000

N = 28

p = < 0,05, r = > 0,334

p = < 0,01, r = > 0,407

TABLA 3

Análisis factorial. Saturaciones de los factores rotados y ordenados y proporción de la varianza explicada por cada uno de ellos procedente de las evaluaciones de los jueces

	Factor 1	Factor 2	Factor 3
DISFONIA 2	0,927	-0,267	0,000
DISFONIA 1	0,904	-0,265	0,000
DISFONIA 3	0,904	-0,257	0,000
VELADA 3	0,839	0,000	-0,442
RONQUERA 3	0,811	-0,481	0,000
RONQUERA 2	0,809	-0,524	0,000
RONQUERA 1	0,764	-0,567	0,000
VELADA 2	0,720	-0,335	-0,450
VELADA 1	0,708	-0,272	-0,525
ÁSPERA 3	0,670	-0,669	0,250
ÁSPERA 2	0,652	-0,631	0,328
TONO 2	-0,303	0,839	0,419
TONO 3	-0,321	0,830	0,411
TONO 1	-0,390	0,790	0,426
ÁSPERA 1	0,620	-0,716	0,000
INTENSIDAD 2	0,000	0,000	0,957
INTENSIDAD 1	0,000	0,000	0,938
INTENSIDAD 3	0,000	0,000	0,927
VARIANZA			
EXPLICADA	12,2069	3,247365	1,2283386
PROP. CUMULATIVA DE LA VARIANZA			
TOTAL	0,678162	0,858571	0,926815

(Las saturaciones inferiores a 0,250 aparecen con el valor 0,000).

y está integrado por las variables disfonía, velada y ronquera de todas las frases, y, en menor medida, por la variable «voz áspera», que también tiene saturaciones altas en el factor 2.

Las variables de tono de las tres frases puntúan alto en el factor 2, así como las variables de «voz áspera» en cada una de las frases, en este caso en sentido negativo. El factor 3 aparece muy claramente como un factor de intensidad, e independiente de los factores 1 y 2.

Podemos, entonces, hablar de tres factores que se delimitan con bastante claridad: El factor 1, o factor «cualidad», el factor 2 o «tono», incluyéndose en este último la voz áspera, o «tipo de voz que se emite en tono muy grave» (Fairbanks, 1960) y el factor 3 o factor de «intensidad».

Discusión

Los resultados de la investigación realizada permiten afirmar que existen parámetros perceptivos —criterios de los jueces— que diferencian distintos tipos de voces, concretamente en este caso voces disfónicas y voces no disfónicas. Los jueces han

diferenciado significativamente a los sujetos disfónicos y no disfónicos en todas las variables de cada frase, excepto en la intensidad de la frase primera. En general, los sujetos del grupo control puntúan más alto en las escalas de intensidad y tono, y más bajo en las escalas de disfonía, velada, ronca y áspera que los sujetos del grupo de disfónicos.

El análisis factorial ha permitido aislar tres factores, en los que saturan de forma diferente las distintas variables. Tales factores se pueden denominar como Factor CUALIDAD —que explica la mayor parte de la varianza total—, Factor TONO, que incluye un tipo de cualidad denominada como «voz áspera» y Factor INTENSIDAD. Los parámetros vocales con claro correlato fisiológico y acústico han sido bien diferenciados por los jueces participantes en el estudio. Asimismo, los jueces han captado aspectos claramente diferenciales en voces disfónicas y no disfónicas centrados en componentes de cualidad, de intensidad y de tono de voz. Todos los parámetros mencionados presentan una independencia relativa con respecto a los otros, llegando a constituir factores independientes, o dimensiones conductuales independientes de la conducta vocal, como sugirió Perkins (1971).

Los sujetos disfónicos participantes en el estudio hablan en un tono más grave, emiten su voz con menos intensidad y manifiestan una cualidad de voz menos «agradable» que sus compañeras no disfónicas. Estas diferencias pueden derivar de un descenso en el rendimiento vocal que afecta a todos los parámetros.

Es muy probable que las diferencias encontradas en la variable intensidad puedan ser debidas a las condiciones propias del sistema de grabación, como volumen de salida, distancia de micro a la boca, nivel de alimentación, etc., aunque, debido precisamente a estas probables influencias ajenas a la experimentación, se intentó controlarlas al máximo antes de iniciar cada grabación. Pero, aun asumiendo estas influencias, éstas no pueden explicar las diferencias encontradas entre los dos grupos.

Por otro lado, parece estar suficientemente comprobado que en voces disfónicas los períodos de fonación son más cortos que en voces no disfónicas (Gould, 1975; Hirano y col., 1980; Eckel y Boone, 1981), siendo esta medida un indicador de la acción de la válvula laríngea y del grado de compresión medial (Daniloff y col, 1980) que a su vez son determinantes directos de la intensidad de la voz. En esta situación, el descenso de la intensidad de la voz de pacientes disfónicos puede quedar bien explicado.

Los resultados respecto a la variable Tono indican claramente que el tono de voz detectado perceptivamente de los sujetos disfónicos es más grave que el de los controles. De acuerdo con Cooper (1973) y Fisher y Logeman (1970), y siguiendo el razonamiento de Boone (1971), parece que el objetivo terapéutico de elevación del tono de voz en pacientes disfónicos puede quedar justificado.

La incógnita que aún perdura al concluir este trabajo es lo que los jueces del estudio han entendido por tono de voz, y si se corresponde exactamente

con la frecuencia fundamental (Fo) o correlato acústico del tono. Lo cierto es que, según los datos obtenidos, el tono de voz puede estar muy relacionado con la cualidad de la misma. De hecho, una de las variables relacionadas con la cualidad de voz o voz «áspera» presenta saturaciones negativas en el factor que hemos aislado e identificado como factor «tono». Puede que para un oyente un tono grave no sea únicamente el que se emite en una frecuencia fundamental menor, sino el característico de voces más duras y más estridentes. Teóricamente, estas variables debieran ser independientes. Este desacuerdo puede hacer pensar en que el oído humano, a pesar de la precisión de sus juicios, no sea capaz de realizar un análisis de la onda compleja por él percibida, lo que justifica la necesidad de utilizar instrumentación acústica para el análisis y el tratamiento de los problemas relacionados con la voz.

Referencias

- Aronson, A. E. (1985): *Clinical Voice Disorders: An Interdisciplinary Approach*. Ed. 2. New York, Thieme, Inc.
- Atkinson, J. E. (1976): Inter e intra-speaker variability in fundamental voice frequency, *J. Acoust. Soc. Am.*, 60, 440-445.
- Boone, D. R. (1971): *The Voice and Voice Therapy*, Englewood Cliffs, New Jersey, Prentice-Hall, Inc.
- Boone, D. R. (1983): *The Voice and Voice Therapy*, Ed. 3. Englewood Cliffs, New Jersey, Prentice-Hall, Inc.
- Burkowsky, M. R., y Vitali, G.J. (1987): *Contemporary Voice Therapy: Children and Adults*, New York, Slosson Educational Publications, Inc.
- Burns, K. L., y Beier, E. G. (1973): Significance of vocal and visual channels in the decoding of emotional meaning, *J. Commun.*, 23, 118-130.
- Cooper, M. (1973): *Modern Techniques of Vocal Rehabilitation*, Springfield, IL, Charles C. Thomas.
- Daniloff, R.; Schuckers, G., y Feth, L. (1980): *The Physiology of Speech and Hearing*, Englewood Cliffs, New Jersey, Prentice-Hall.
- Davitz, J. R. (1964): *The Communication of Emotional Meaning*, New York, Mc-Graw-Hill.
- Dixon, W. J. (ed.) (1983): *BMDP Statistical Software*, California, University of California Press.
- Eckel, F. C., y Boone, D. R. (1981): The s/z ratio as an indicator of laryngeal pathology, *J. Speech Hear. Disord*, 46, 147-149.
- Esteve, J. M.: Enfermedades de larga duración de los docentes, *Comunidad Escolar*, 18 enero.
- Fairbanks, G. (1960): *Voice and Articulation Drillbook*, New York, Harper and Row.
- Fairbanks, G., y Hoaglin, L. W. (1941): An experimental study of the durational characteristics of the voice during the expression of emotion, *Speech Monographs*, 8, 85-90.
- Fierro, A. (1983): Diferencial Semántico. En Ballesteros Fernández: *Psicodidáctico. Apéndice. Cuaderno de Prácticas*, UNED.
- Fisher, H. B., y Logemann, J. (1970): Objective evaluation of therapy for vocal nodules: A case report. *J. Speech Hear. Disord*, 35, 277-285.
- Gólderos, A.; Martínez, R.; Nombela, J. R.; Pardo, M.; Santos, J., y Muñoz, E. (1980): Comunicación hombre-máquina por voz (I). Modelo de producción de la voz, *Mundo Electrónico*, 96, 47-50.
- Gould, W. J. (1975): Quantitative assessment of voice function in microlaryngology, *Folia Foniatr.*, 27, 190-220.
- Hirano, M.; Gould, W. J.; Lambiase, A., y Kakita, Y. (1980): Movements of selected points on a vocal fold during vibration, *Folia Phoniatr.*, 32, 39-50.
- Hollien, H., y Moore, G. P. (1966): Measurements of the vocal fold during changes in pitch, *J. Speech Hear. Res.*, 12, 616-628.
- Jiménez, A. (1985): *Marcadores emocionales en la conducta vocal*, Universidad Autónoma de Madrid, Tesis Doctoral.
- Kramer, E. (1963): Judgment of personal characteristics and emotions from nonverbal properties of speech, *Psychol. Bulletin*, 60, 408-420.
- Laver, J. (1980): *The Phonetic Description of Voice Quality*, Cambridge, Cambridge University Press.
- Michel, J. F., y Wendahl, R. (1971): Correlates in voice production. En L. E. Travis (ed.): *Handbook of Speech Pathology and Audiology*, New York, Appleton-Century-Crofts.
- Moncur, J. P., y Brackett, I. P. (1974): *Modifying Vocal Behavior*, New York; Harper & Row.
- Osgood, Ch.; Suci, G. J., y Tannebaum, P. H. (1957): *The Measurement of Meaning*, University of Illinois Press, Urbana IL (traducción española: *La Medida del Significado*, Madrid, Gredos, 1976).
- Perkins, W. H. (1971): Vocal function: Assessment and therapy. En L. E. Travis (ed.): *Handbook of Speech Pathology and Audiology*, 505-534, Englewood Cliffs, New Jersey, Prentice-Hall, Inc.
- Quilis, A. (1981): *Fonética Acústica de la Lengua Española*, Madrid, Gredos.
- Reed, Ch, G. (1980): Voice therapy: A need for research, *J. Speech Hear. Disord*, 45, 143-156.
- Rosenthal, R. (1982): Conducting judgment studies. En K. R. Scherer y P., Ekman (eds.): *Handbook of Methods in Nonverbal Behavior Research*, Cambridge, Cambridge University Press.
- Scherer, K. R. (1979): Personality markers in speech. En K. R. Scherer and H. Giles (eds.): *Social Markers in Speech*, 147-202, Cambridge, Cambridge University Press.
- Scherer, K. R. (1982): Methods of research on vocal communication: Paradigms and parameters. En K. R. Scherer, y P. Ekman (eds.): *Handbook of Methods in Nonverbal Behavior Research*, Cambridge, Cambridge University Press.
- Scherer, K. R.; Koivumaki, J., y Rosenthal, R. (1972): Minimal cues in the vocal communication of affect: Judging emotions from contentmasked speech, *J. Psycholing. Research*, 1, 269-285.
- Schultz-Coulon, H. J. (1975): On the measurement and the evaluation of the individual median frequency of the speaking voice. And experimental study, *Folia Phoniatr.*, 17, 209-214.
- Sogon, S. A. (1975): A study of the personality factor which affects the judgment of vocally expressed emotions, *Japan. J. Psychol.*, 46, 247-254.
- Soskin, W. F., y Kauffman, P. E. (1961): Judgement of emotions in word-free voice samples, *J. Comun.*, 11, 73-81.
- Wedin, S.; Leanderson, R., y Wedin, L. (1978): Evaluation of voice training. Spectral analysis compared with listener's judgements, *Folia Phoniatr.*, 30, 103-112.
- Williams, C. E., y Stevens, K. N. (1969): Emotions and speech: Some acoustical correlates, *J. Acoust. Soc. Am.*, 52, 1238-1250.
- Wilson, F. B., y Broda-Masiuk, D. (1987): *Voice Disorders*, Voice Tapes, Bellingham, VA.
- Zuckerman, M.; Lipets, M. S.; Koivumaki, J. H., y Rosenthal, R. (1975): Encoding and decoding nonverbal cues of emotion, *J. Personal. Social Psychol.*, 32, 1068-1076.