

La respuesta emocional de ansiedad de los experimentos: un estudio mediante parámetros vocales

I. MARTÍN LÓPEZ y J. A. TALAVERA

Universidad Autónoma de Madrid



Resumen

Hemos pretendido estudiar la evolución que puede tener la respuesta emocional de ansiedad en el curso de los experimentos con sujetos humanos. Para ello, diseñamos un esquema de dos tareas que exigían un mismo tipo de respuesta verbal pero eran diferentes en cuanto a su complejidad cognitiva. Fueron aplicadas ambas tareas, pero con el orden balanceado, a dos grupos con un total de 20 sujetos normales. El estado de ansiedad, o estrés emocional, fue determinado mediante técnicas digitales de análisis de voz en la parte inicial y final de las tareas. En contra de nuestra hipótesis de partida, inicialmente los niveles de ansiedad no tuvieron diferencias significativas y, por tanto, no estaban relacionados con la dificultad de la tarea. Después, estos niveles siguieron cursos diversos, incrementándose en el caso de la ejecución más difícil. Esto indica que los sujetos posteriormente son capaces de controlar y ajustar sus estados de ansiedad de acuerdo con las exigencias de las tareas que están realizando. También hemos encontrado en los sujetos que tienen una personalidad más ansiosa (según las puntuaciones en el cuestionario STAI en su forma de rasgo) que su principal diferencia no estriba en que alcanzan inicialmente mayores niveles de ansiedad de estado si no que reducen en menor proporción o mantienen, respecto a los valores iniciales, sus niveles de ansiedad durante la ejecución de las tareas sencillas.

Palabras clave: emoción, ansiedad, conducta vocal.

Subjects' anxiety in experimental situations: A study of their vocal parameters

Abstract

The aim of this study was human subjects' anxiety in experimental situations. Digital analysis of vocal utterances allowed us to study this phenomenon through all the steps of an experimental situation. Two tasks were designed. They had different levels of cognitive complexity but both required the same kind of verbal response. Two groups of 20 subjects had to complete the tasks in different order. A digital analysis of subjects' voice was performed along the task. Results show that levels of anxiety were the same at the beginning of both tasks (the ease and the difficult one) but increased while subjects were performing the difficult task. Therefore, subjects seem to control and adjust their anxiety according to the kind of task they were performing. Results also pointed out that there were not initial differences in anxiety according to subject's level in trait-anxiety. However individuals with highest trait-anxiety rates were more inefficient than others to reduce their initial levels of anxiety as state.

Keywords: emotion, anxiety, vocal behavior.

Agradecimientos: La investigación ha sido en parte financiada por una beca del CAICYT proyecto PB85-226. Queremos agradecer a José Miguel Fernández Dols su indispensable apoyo técnico y moral, así como, la colaboración de los sujetos experimentales.

Dirección del autor: I. Martín López, Facultad de Psicología (despacho 43) Universidad Autónoma de Madrid, 28049 Madrid.

INTRODUCCION

Sabemos que los sujetos, ante muchas situaciones normales de la vida, reaccionan con una respuesta de ansiedad o de estrés psicológico. Es una respuesta compleja elicitada por estímulos inespecíficos, según muchos autores, que puede ser estudiada desde una gran variedad de perspectivas y que está recibiendo una atención creciente por parte de la medicina actual dada la influencia que, se considera, tiene en la salud y en la enfermedad (Weiner, 1989). A nosotros sólo nos ha interesado en este estudio su vertiente emocional que a su vez conjuga componentes psicológicos y fisiológicos. Además, antes de poder generalizar a situaciones de la vida normal, queremos conocer las respuestas que se dan en situaciones experimentales. Un campo muy abonado para ellas dado que en un gran porcentaje de estudios empíricos los sujetos perciben las situaciones experimentales como artificiales, desconocidas y, por tanto, posiblemente generadoras de ansiedad.

Desgraciadamente en muy pocos trabajos experimentales se intenta controlar dicha variable. Incluso en trabajos dedicados especialmente en su análisis no se conoce adecuadamente sus niveles. Típicamente, en dichos estudios se compara la eficacia de un grupo de sujetos cuando realizan unas tareas en una condición de alta ansiedad respecto a otro, o el mismo, grupo de sujetos realizando las anteriores tareas en condición de baja ansiedad. Si las tareas, por sí mismas, indujeran importantes niveles de ansiedad comparado con las condiciones «ansiógenas», como puede ser el caso del experimento, por ejemplo, de Shapiro y Lim (1989) en el cual la ejecución de las tareas era medida por tiempos de reacción a estímulos visuales y las condiciones de alta y baja ansiedad se obtenían con música «relajante» o «estresante», entonces las diferencias encontradas podrían ser debidas a las características particulares de los estímulos inductores de ansiedad más que las diferencias en los niveles de ansiedad inducidos en los sujetos. En otros experimentos, aunque se mide adecuadamente el nivel de ansiedad con parámetros fisiológicos junto con escalas subjetivas, se pasa por alto el carácter transitorio y rápidamente cambiante de la ansiedad y lo que medimos hace algunos minutos es posible que ahora ya no valga. Para ilustrar este error nos puede servir de referencia el experimento de Foa, McNally y Murdock (1989) en el que se controlaban bien los niveles de ansiedad con la tasa cardíaca y escalas autoaplicadas, pero estas medidas eran tomadas justo después de haberles inducido verbalmente las condiciones ansiógenas y algunos minutos antes de la realización de las tareas. Simplemente este hecho podría dar una explicación a las diferencias no significativas que encontraron los autores: cuando los sujetos realizaban las tareas ya habría desaparecido la ansiedad inducida.

En el presente trabajo nos hemos propuesto dar respuesta a los anteriores problemas y analizar cuantitativamente, por un lado, cómo las tareas que están ejecutando los sujetos influyen en sus niveles de ansiedad. En este aspecto, los resultados experimentales parecen indicar que la respuesta emocional de ansiedad es mayor con tareas difíciles que con fáciles (Brenner, Shipp y Morrissey, 1983; Eysenk 1982; Weiner y Schneider 1971), tal como el sentido común sugiere. Por otro lado, sin embargo, pocos estudios, han observado cómo evoluciona esa respuesta de ansiedad. Por tanto,

no sólo deseamos saber qué valores toma sino cómo evoluciona a lo largo del experimento. ¿La ansiedad aumenta, disminuye o se mantiene constante?

Para la determinación cuantitativa e instantánea de los niveles de ansiedad nos hemos basado en el lenguaje de los sujetos, en concreto en la realización de un análisis instrumental de los parámetros acústicos de sus respuestas vocales. Numerosos estudios han puesto de manifiesto la importante influencia de la ansiedad, tanto en la parte verbal como en la parte no verbal, sobre las producciones vocales. Así, vemos que en situaciones de ansiedad los sujetos hablan con menos fluidez, repiten más las palabras y éstas, a su vez, son de menor longitud (Meulemans, 1974). Suelen emplear un lenguaje ritualizado con una abundante profusión de palabras sin contenido como «entonces», «de hecho» (Lalljee y Cook, 1975, Gottschalk y Frank, 1967) con tartamudeos y numerosas exclamaciones no léxicas del estilo de «ah» (Musumeci, 1975, Ragsdale, 1977). Dentro del contexto de una comunicación entre dos personas, cuando están en estados de ansiedad moderadamente alta sus conversaciones tienen menos pausas, risas, interrupciones y expresiones verbales formuladas como preguntas (Hurley, 1983).

Incluso si prescindimos de la perspectiva comunicativa y de la parte lingüística del lenguaje y nos centramos en el estudio de las cualidades de las voces de los sujetos en estados de ansiedad, encontramos que los parámetros extraídos del análisis instrumental de la onda acústica también son sensibles a los niveles de ansiedad como veremos más adelante. En general, tanto la ansiedad como las emociones específicas parecen tener una importante influencia sobre todos los aspectos de la producción vocal y esto a nivel neurofisiológico se explica por las conexiones que existen entre las áreas cerebrales del lenguaje y el sistema límbico (Jürgens, 1983).

Es posible identificar en algunos parámetros vocales tendencias que se dan en la mayoría de los sujetos. Por ejemplo, la frecuencia fundamental media, término científico relacionado con el tono de voz y simbolizado por F_0 (en adelante en el artículo), aumenta con la ansiedad o el estrés psicológico (Williams y Stevens, 1972). También lo hace la intensidad (Friedhoff, Alpert y Kurtzberg, 1966, Menahem, 1983) y la proporción de energía espectral en el rango aproximado del primer formante (Simonov y Frollov, 1973). Otras características de las respuestas verbales que son sensibles a la ansiedad son las temporales (Hicks, 1979) y los parámetros relacionados con la F_0 como, por ejemplo, el suelo del contorno de tono (Tolkmitt y Scherer, 1986). Sin embargo, uno de los mayores problemas con que se encuentran los estudios sobre la ansiedad es que existen claras diferencias individuales. No todos los sujetos responden, tanto fisiológica como psicológicamente, de la misma forma (Eysenck, 1983). Esto mismo se ha observado que ocurre cuando tratamos de cuantificar el estado de ansiedad con los parámetros vocales (Almeida, Fleischmann, Heike y Thormann, 1975, Hecker, Stevens, Bismarck y Williams, 1968, Scherer, 1986).

Por último, uno de los parámetros vocales más utilizados ha sido la atenuación que parecen producir los sujetos cuando están en situaciones de ansiedad sobre unas pequeñas oscilaciones superpuestas a la onda acústica en el entorno de los 10 Hz. Estas oscilaciones se consideran supuestamente debidas a los minúsculos temblores que presentan todos los grandes mús-

culos y probablemente también los fonatorios. Aunque para la detección de mentiras, donde se han realizado multitud de experimentos (Barland, 1973), podemos poner en duda su validez (Hollien, 1980, Brenner, Branscomb y Schwartz, 1979), sí es generalmente admitido que la reducción de los pequeños temblores de frecuencia baja, en la onda vocal, correlacionan significativamente con los estados de ansiedad (Inbar, Eden y Kaplan, 1977, Smith, 1977, Brockway, 1979, Brenner, Shipp y Morrissey, 1983).

Podemos resumir, por tanto, nuestra hipótesis esperando encontrar mayores niveles de ansiedad durante toda la ejecución de la tarea más compleja, lo que se verá reflejado en una menor energía en el entorno de los 10 Hz, así como, mayor frecuencia fundamental, intensidad y energía de los primeros formantes, entre otros parámetros de las respuestas vocales, en comparación con esos mismos parámetros de las respuestas durante la ejecución de la tarea sencilla.

METODO

Sujetos

Un total de 20 sujetos adultos, con edades comprendidas entre 25 y 57 años, nativos hispano parlantes, sin historiales de anormalidades neurofisiológicas, problemas de lenguaje o de audición, fueron asignados a dos grupos: A y B. La subdivisión del total de sujetos en dos grupos tuvo el objetivo de controlar el posible efecto del orden de presentación de las tareas. A cada grupo se presentó las mismas tareas pero en orden diferente. La asignación de los sujetos a los grupos fue realizada buscando la máxima homogeneidad entre ambos. Así, cada grupo contó con el mismo número de mujeres que de hombres y se buscó que la edad y el nivel cultural medio fueran similares, tal como se aprecia en la tabla I.

Como datos adicionales se tomaron: el nivel cultural, operacionalizado en cuatro categorías que van desde uno para estudios elementales hasta cuatro para estudios universitarios. El estatus socioeconómico, categorizado en cinco niveles con puntuaciones de uno a cinco: muy bajo, bajo, medio, alto y muy alto. Además se evaluaron, de forma simplificada, la cantidad de interacciones íntimas al mismo nivel de la jerarquía familiar, tanto del pasado, mediante el número de hermanos y de su posición cronológica, como del presente, mediante el índice que nosotros denominamos «compañía». El número de hermanos lo hemos categorizado en tres niveles con puntuaciones de uno a tres: ningún hermano, un hermano, más de un hermano. El índice «posición dentro de la estructura familiar», fue operacionalizado en tres niveles con puntuaciones del uno al tres: primogénito, posición intermedia, última posición. Por último, el índice que denominamos «compañía» se categorizó en dos niveles: uno si vivían solos y dos si estaban casados o compartían la vivienda con más personas.

Estímulos: protocolos grabados

Sobre una cinta magnética de bobina abierta, se grabaron dos tareas experimentales. Antes y después de cada tarea experimental se introdujeron intervalos de sombreado binaural para independizar las tareas entre sí y es-

TABLA I

Características de los dos grupos experimentales y del total de sujetos. Figuran en la tabla la proporción de varones y hembras (V/H), medias y desviaciones típicas de la edad y de la ansiedad de rasgo medida mediante la escala STAI-R. Complementan la tabla cinco índices sociológicos: nivel cultural, estatus socioeconómico, número de hermanos, posición dentro de la estructura familiar y compañía

Grupo	Proporc. V/H	Edad		Ansiedad		N. cultural		E. Socioec.		N.º hermanos		Posición		Compañía	
		μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
A	5/5	38	12	16,2	7,7	3,7	0,46	3,1	0,73	2,7	0,47	1,6	0,84	1,7	0,48
B	5/5	36	9,1	17,7	4,8	3,8	0,40	3	0,81	2,7	0,48	2	0,94	1,9	0,31
Total	10/10	37	10,5	16,9	6,3	3,75	0,43	3,05	0,75	2,7	0,48	1,8	0,89	1,8	0,41

tablecer una línea base. La primera tarea experimental consistía en la presentación, a intervalos regulares de 30 segundos, de palabras comunes de ocho letras. Se pedía al sujeto que, a su ritmo, las deletreara en orden inverso. En los cinco minutos que duraba esta tarea, se habían presentado un total de 10 palabras. La segunda tarea experimental consistía en la presentación de la vocal castellana /a/ con una cadencia de tiempo aleatoria con distribución discreta y homogénea en un rango de 5 ± 3 segundos. En los cinco minutos de duración de la prueba se habían presentado un total de 60 vocales. A los sujetos se les pedía que repitieran lo más rápido posible, el mismo fonema que oían. Los intervalos de sombreado biaural (repetición en voz alta de lo que se va escuchando por ambos oídos) eran de un minuto de duración y fueron extraídos de «El Camino», novela de Miguel Delibes, la cual posee un lenguaje especialmente sencillo. Los tres intervalos fueron simplificados, homogeneizados, igualados en longitud y narrados con una cadencia de 120 palabras/minuto.

A partir de la cinta maestra se formaron dos protocolos con exactamente el mismo contenido pero con una posición diferente de las tareas experimentales tal como se esquematiza a continuación:

Protocolo A B1 INV B2 DIR B3
Protocolo B B1 DIR B2 INV B3

- B1 B2 B3: intervalos de neutralización o línea base; sombreado biaural a 120 palabras/minuto.
- INV: tarea experimental de deletreo inverso de palabras de ocho letras.
- DIR: tarea experimental de repetición directa de la vocal castellana /a/.

Entre cada tarea había una pausa de descanso de un minuto de duración que se aprovechaba para incluir las instrucciones del siguiente intervalo. Estas instrucciones eran narradas con voz masculina mientras las tareas con voz femenina. Cada protocolo completo duraba un total de aproximadamente 18 minutos.

Procedimiento experimental y de análisis

Las pruebas eran pasadas a los sujetos de forma individual y estaban formadas por los protocolos grabados en cinta magnética, escalas autoaplicadas, entre las que se encontraba la escala STAI (Spielberger, Gorsuch, y Lushene, 1970) en su forma de rasgo, y una entrevista biográfica a través de la cual podíamos determinar las características individuales y sociológicas de los sujetos. El protocolo grabado A fue pasado a los sujetos del grupo A mientras que el protocolo B a los sujetos del grupo B.

A través de unos auriculares, los sujetos escucharon las instrucciones y recibieron los estímulos. Sus respuestas fueron grabadas mediante un micrófono situado a siete centímetros de los labios y formando con la horizontal un ángulo hacia arriba de cuarenta y cinco grados que evitaba los ruidos producidos por las turbulencias de aire exhalado al incidir sobre la superficie activa del micrófono. La posición y distancia de los labios al micrófono se mantuvieron constantes aunque los sujetos movieran la cabeza, por estar este último unido fijamente a los auriculares. De esta forma se pudo medir exactamente la intensidad vocal sin distorsiones provocadas por variaciones de posición o distancia al micrófono. Además de las respuestas de los sujetos, por el otro canal, el derecho, se grabaron simultáneamente los estímulos presentados. Esto permitió una determinación exacta de los tiempos de reacción.

Las cintas grabadas, con respuestas y estímulos, fueron digitalizadas para realizar un análisis posterior con ordenador. Esta digitalización abarcó dos formas: de un canal y de dos canales. En la de un canal, que constituyó la principal, los intervalos de muestreo fueron de un segundo de duración a una frecuencia de 8 kHz y una resolución de 10 bits. Estos intervalos recogían la primera y la última vocal /a/ de cada tarea. Así, se contó con un total de cuatro muestras, de un canal, por cada sujeto. Se eligió, como estímulo a reproducir, la vocal /a/ por que corresponde con la posición neutra de las estructuras fonoarticulatorias del tracto vocal humano y, por tanto, menos susceptible de variar en medidas repetidas.

En la digitalización para dos canales, dado que su objetivo era sólo el análisis del tiempo de reacción que transcurría desde la producción de la señal (vocal /a/ emitida por el experimentador y grabada en el canal derecho) hasta la emisión de respuesta del sujeto (repetición de vocal /a/ y grabada en el canal izquierdo), las frecuencias de muestreo fueron bajadas a 200 Hz, con una resolución de 12 bits y duraciones mucho mayores, de 30 segundos. Este método sólo fue utilizado en la repetición directa de las vocales pues, en la tarea de deletreo inverso, el sistema automático tenía un índice alto de fallos, dado que los tiempos eran extremadamente largos y muchos sujetos intercalaban exclamaciones antes de las respuestas correctas. Al ser la duración de estos tiempos suficientemente prolongada pudieron medirse fiablemente con cronómetro.

Los parámetros calculados por los programas de ordenador pueden clasificarse en las siguientes cuatro categorías:

- a) Frecuencia fundamental y parámetros asociados.
- b) Parámetros temporales.
- c) Parámetros de intensidad.
- d) Distribución espectral.

a) Dentro de la primera categoría, el sistema automático de análisis determinó la media estadística, desviación típica, mediana y moda de los valores inversos de cada período fundamental hallado en las fonaciones de los intervalos de muestreo. La media estadística, así hallada, es lo que denominamos frecuencia fundamental media (F_0). También se determinaron los percentiles del 5 % (suelo de la F_0) y del 95 % (techo de la F_0). La frecuencia fundamental, contrariamente a lo que sucede con los instrumentos musicales, exhibe unas variaciones de un período a otro relacionadas con la fisiología del aparato fonatorio y, posiblemente, del estado psicológico del individuo. Se evaluaron estas variaciones hallando las diferencias entre períodos fundamentales consecutivos y calculando su porcentaje respecto al período medio. El sistema automático calculó estos valores y determinó su media estadística, desviación típica, valores máximos y mínimos. Finalmente, dentro de esta categoría, se hallaron las roturas de tono: desapariciones transitorias de la onda fundamental, en tanto por ciento.

b) Como parámetros temporales, se determinaron las medias de los tiempos de fonación e intervalos de pausa. Dentro de las fonaciones, se desglosaron los intervalos sordos (cuando aparece energía vocal pero no se detecta onda fundamental) de los intervalos sonoros (cuando existe tanto energía como onda fundamental). Estas últimas magnitudes son expresadas también en tanto por ciento. Además de estos parámetros temporales clásicos dentro del análisis de la voz, el sistema determinó, utilizando los dos canales, una medida también clásica, pero en este caso de los estudios de atención: tiempos de reacción.

c) Dentro de la tercera categoría, se consideran dos tipos de intensidad: una instantánea y otra promediada. La instantánea es el valor máximo que alcanza la onda dentro de un período fundamental y la promediada es la integración de todos los valores rectificadas a lo largo de cada fonación. Mientras que de la intensidad promediada se determinaron las medias, desviaciones típicas, valores máximos y valores mínimos dentro de cada intervalo de análisis, de la instantánea determinamos esos mismos estadísticos pero de su variabilidad entre períodos contiguos. Esta variabilidad tiene el mismo algoritmo de cálculo que la variabilidad ciclo a ciclo de la frecuencia y la denominamos, por tanto, variabilidad ciclo a ciclo de la amplitud.

d) Dentro de la última categoría se determinaron las energías de cinco bandas de frecuencia: 8-16 Hz, 24-32 Hz, 240-440 Hz, 440-1.000 Hz, 1-3 kHz. La energía en cada banda fue expresada en valor porcentual respecto a la suma de las energías de todas las bandas. Además, se calcularon la frecuencia en hercios y energía en decibelios de los tres primeros formantes.

Instrumentación

La cinta maestra se hizo con el magnetófono de bobina abierta TANDBERG modelo TD20A y posteriormente editada sobre una cassette en doble cara mediante la pletina AKAI LS705D.

Las grabaciones de los sujetos se hicieron con una doble pletina Marantz, modelo SD275, con los siguientes ajustes: Amplificación de entrada: 6 para el canal izquierdo y 4 para el canal derecho, reductor de ruido cancelado y cinta de óxido de cromo. En la entrada de la pista izquierda se

conectó el micrófono AKAI ADM80. A la entrada derecha se conectó la salida del aparato reproductor. Para algunos sujetos se utilizó la pletina AKAI LS705D convenientemente calibrada para obtener el mismo nivel de grabación (4,5 canal izquierdo y 7,5 canal derecho) que la anterior. La reproducción de los protocolos que escuchaban los sujetos se realizó en un cassette portátil Marantz CP430 a un nivel de reproducción de 1/4 del nominal, sin reducción de ruidos y salida por auriculares SOUND SA1322.

En el análisis instrumental se reprodujeron las cintas grabadas con la pletina Marantz SD275. Su salida del canal izquierdo estaba conectada a un filtro Cibertec LH2 calibrado con ganancia unidad, modo pasabanda, con frecuencia superior de corte de 3 kHz y frecuencia inferior de corte de 2 Hz. La salida del filtro iba a una tarjeta de adquisición, por su entrada directa, de Software Research Corporation; Micro Speech Lab serie 89 3.0241. La tarjeta estaba instalada en un ordenador Tandom Plus. En dicho ordenador se hicieron las operaciones de muestreo digital con el paquete de software Micro Speech Lab y el análisis de los parámetros vocales con un paquete de programas desarrollados en el ICAI, Universidad Pontificia de Comillas.

RESULTADOS

Diferencias intertareas

Los resultados en los tiempos de reacción nos confirman nuestra suposición inicial de que el deletreo inverso es una tarea mucho más compleja cognitivamente que la repetición monovocálica. Mientras que los sujetos tardaron una media de 3,143 segundos en pronunciar la primera letra correcta de la primera palabra y 3,480 segundos en la última, en la repetición monovocálica tardaron tan sólo 399 milisegundos y 319 milisegundos respectivamente. Es decir, los sujetos fueron unas diez veces más rápidos en responder en una tarea experimental que en otra ($p < 0,001$, $F_{1,19} = 101,57$, ANOVA multivariable para medidas repetidas). Esta diferencia resultó independiente del orden de presentación de las tareas, es decir, del protocolo realizado por los sujetos ($p > 0,1$).

Aunque las respuestas verbales que se comparan son las mismas en todos los casos, vocales /a/, las exigencias de las tareas imponen características diferenciadoras. Así, vemos en la tabla II (columna INI) que ya inicialmente hay diferencias entre los parámetros vocales. Entre las que se consideran estadísticamente significativas están las relativas a los tiempos de fonación (t-Student, $N = 20$, $p < 0,001$). Es curioso que, por ejemplo, cuando los sujetos inicialmente repetían las vocales lo hicieran pronunciándolas de manera más corta que cuando las deletreaban en orden inverso. En particular, los intervalos sonoros, parte de la fonación que contiene onda fundamental, fueron mucho menores ($p < 0,001$). Por otro lado, otro parámetro no relacionado con los tiempos, la desviación típica de la variabilidad ciclo a ciclo de la intensidad máxima de la onda fundamental, tuvo inicialmente mayor magnitud ($p = 0,032$) en las respuestas de la tarea de repetición vocálica que en las de deletreo inverso.

Según van siendo ejecutadas, las tareas van imprimiendo otras características particulares a las voces. Mientras los parámetros relativos a tiempos

de fonación se van, por un lado, igualando, por otro, van apareciendo nuevas diferencias significativas en más variados y abundantes parámetros. Así, los valores medios de las frecuencias fundamentales son mayores en las vocales deletreadas que en las repetidas (t-Student, $N = 20$, $p = 0,012$) en los intervalos finales. También lo son algunos de los estadísticos relacionados con la frecuencia fundamental como las modas ($p = 0,013$), las medianas ($p = 0,003$) y los techos —percentiles del 95 % ($p = 0,033$). Por otro lado, se acentúan las diferencias en los parámetros relativos a la variabilidad de la intensidad máxima de cada período. Así, las vocales repetidas alcanzaron una mayor variabilidad ciclo a ciclo media de la intensidad que las vocales /a/ deletreadas ($p = 0,044$). Esto mismo ocurrió con sus desviaciones típicas ($p = 0,027$) y los valores máximos y mínimos ($p = 0,036$, $p = 0,028$). Además, las vocales /a/ deletreadas en sentido inverso tuvieron menos energía en la banda de frecuencias bajas (24-32 Hz) y más en las cercanas al segundo formante, frecuencias en el entorno de los 1.500 Hz, que las repetidas directamente ($p = 0,022$, $p = 0,011$).

Diferencias intratarea

Si observamos las variaciones, dentro de la misma tarea, que sufren los parámetros vocales en el curso de su realización, encontramos que la repetición monovocálica induce incrementos significativos en los tiempos de fonación ($p = 0,001$) y en los intervalos sonoros ($p < 0,001$) mientras que induce decrementos en la variabilidad ciclo a ciclo de la frecuencia fundamental ($p = 0,005$). El deletreo inverso, por otro lado, produce sensibles incrementos en la intensidad promediada de las fonaciones ($p = 0,02$) y decrementos en la variabilidad ciclo a ciclo mínima de las amplitudes máximas ($p = 0,014$).

DISCUSION

Hasta el momento sólo hemos revisado los parámetros vocales sin ninguna relación con los constructos psicológicos. Ahora vamos a interpretarlos de acuerdo con la respuesta emocional de ansiedad. Al comienzo de la ejecución, las respuestas verbales, en ambas tareas, no tuvieron significativamente diferentes ninguno de los parámetros vocales que se relacionan con la ansiedad, al menos según lo generalmente aceptado, excepto los tiempos medios de fonación, de los cuales hablaremos más adelante. Esto nos permite afirmar, en contra de nuestra hipótesis de partida, que inicialmente los sujetos tienen los mismos niveles de ansiedad independientemente de la dificultad de la tarea que realicen.

En el transcurso de la ejecución de las tareas, ciertos parámetros relacionados con la ansiedad van diferenciándose. Así, observamos al final de una frecuencia fundamental media, mediana y moda mayores en el deletreo inverso que en la repetición monovocálica. Aunque existen diferencias individuales, tal como se discutió en la introducción, la ansiedad suele inducir incrementos de la frecuencia fundamental (Scherer 1979), por tanto, estos datos sugieren que los sujetos están bajo un estado de mayor ansiedad en la parte final de la tarea de deletreo inverso que en la parte final de la

TABLA II

Parámetros acústicos de las vocales castellanas /a/ pronunciadas al comienzo y final de las tareas de repetición monovocálica y de deletreo inverso de palabras de ocho letras. En las columnas de la derecha figuran marcadas las diferencias que han sido significativas entre las vocalizaciones iniciales y finales dentro de cada tarea (R.M. y D.I.) y comparando entre una tarea y otra las fonaciones iniciales (INI.) y finales (FIN.), así como las variaciones (VAR.)

Var.	Denominación	Repe. monovocálica				Deletreo universo				Unidad	Significación estadística				
		Inicial		Final		Inicial		Final			Intratar.		Intertareas		
		μ	σ	μ	σ	μ	σ	μ	σ		R.M	D.I.	INI.	FIN.	VAR.
FUM	Frecuencia fundamental. Media	156	41,8	155	51	162	50,6	170	58	Hz			*	*	
FUD	Frecuencia fundamental. Desviación t.	10,3	11,2	10,2	10	20	17,4	16,1	16	Hz					
FMO	Frecuencia fundamental. Moda	156	44,2	154	50,5	156	47	175	71	Hz			*		
FME	Frecuencia fundamental. Mediana	161	42,9	159	51,1	166	49,8	174	56	Hz			**		
FTE	Frecuencia fundamental. Techo (95 %)	180	53,8	177	62,5	198	73	202	80,4	Hz			*		
FSU	Frecuencia fundamental. Suelo (5 %)	145	39,8	143	48,2	140	43,5	149	49,7	Hz					
VFM	Var. ciclo a ciclo de Fo. Media	2,43	0,80	1,87	0,71	2,05	0,61	1,93	0,62	%	**				
VFD	Var. ciclo a ciclo. Desviación t.	2,15	0,64	1,95	0,75	2,09	0,71	1,97	0,78	%					
VFMAX	Var. ciclo a ciclo. Máxima	7,21	2,50	7,06	3,20	7,98	3,11	7,65	3,73	%					
VFMIN	Var. ciclo a ciclo. Mínima	0,5		0,0		0,0		0,0		%					
RTO	Roturas de tono	7,83	12	7,33	11,4	6,27	10	3,84	7,59	%					
TMF	Tiempo de fonación	0,17	0,06	0,26	0,09	0,29	0,12	0,28	0,10	Seg.	***		***	*	
ISO	Intervalo sordo. Porc. acumulado	13,7	15,9	13,4	13	15,0	9,90	17,6	17,8	%					
ISN	Intervalo sonoro. Porc. acumulado	13,7	4,92	23,1	8,17	27,5	12,1	31,9	13,5	%	***		***	*	
PPA	Intervalo pausa. Porc. acumulado	64,7	25,4	56,1	20,2	51,1	24,9	46,6	26,1	%			*		
TR	Tiempos de reacción. Media	399	85,2	319	85,2	3143	1196	3480	1523	mSeg.			***	***	
IMF	Intensidad promediada de fonac.	48,6	22,9	48,4	26,6	44,8	22,6	50,5	22,7	—		*		*	
VAM	Var. ciclo a ciclo de I. Media	16,5	6,05	14,7	5,71	13,4	3,94	12,1	5,49	%			*	*	
VAD	Var. ciclo a ciclo. Desviación t.	18,4	7,10	16,3	7,40	14,3	5,68	12,4	5,14	%			*	*	
VAMAX	Var. ciclo a ciclo. Máxima	70,5	24,2	70,8	32,2	68,4	32,1	54,6	21,5	%			*	*	
VAMIN	Var. ciclo a ciclo. Mínima	1,47	1,62	1,36	2,17	1,01	1,23	0,22	0,36	%		*	*		
EP	Porcentaje de energía en 8-16 Hz	2,96	4,04	3,44	4,64	1,81	2,21	1,89	2,64	%					
EPR	Porcentaje de energía en 24-32 Hz	2,44	2,88	3,07	3,72	1,53	2,09	1,42	2,30	%			*		
ES	Porcentaje de energía en 40-440 Hz	18,1	10,8	20,9	12,3	23,9	16,4	24,7	13,7	%					
ET	Porcentaje de energía en 440-1000 Hz	55,9	16,2	51,9	16,8	55,5	17,1	52,4	17,2	%					
EC	Porcentaje de energía en 1000-3000 Hz	23	15,4	23,7	21,1	18,7	12,5	20,8	13,9	%					
FFP	Frecuencia del primer formante	718	206	721	218	705	219	615	248	Hz					
FEP	Energía del primer formante	-22	8,60	-22	11	-25	9,97	-21	10	dB					
FFS	Frecuencia del segundo formante	1519	59,3	1501	56,8	1496	48,8	1508	53,3	Hz					
FES	Energía del segundo formante	-56	13,5	-60	17,4	-58	16,9	-55	16,8	dB			*	*	
FFT	Frecuencia del tercer formante	2230	415	2122	411	2156	373	2256	464	Hz					
FET	Energía del tercer formante	-70	12,6	-73	17,5	-73	13	-70	13	dB					

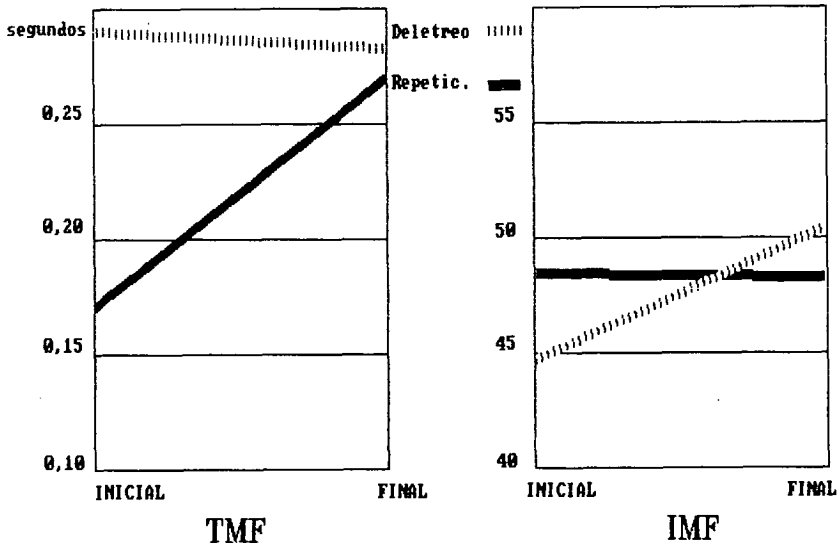
* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

tarea de repetición vocálica. Además, la energía del segundo formante, situado alrededor de los 1.500 Hz, es significativamente mayor, mientras que es menor la energía relativa en la banda de bajas frecuencias 24-32 Hz. Esto parece indicar un cierto desplazamiento de la energía espectral, desde las bandas de baja frecuencia hacia las bandas que contienen los formantes en el curso del deletreo inverso. Hay datos para suponer que un aumento de la ansiedad tiende a producir este mismo desplazamiento de energía espectral (Hecker, Stevens, Bismarck y Williams, 1968, Kuroda, Fujiwara, Okamura y Utsuki, 1976).

Tanto este desplazamiento de la energía espectral, como el incremento de la frecuencia fundamental según se va deletreando en modo inverso, expresan, con bastante contundencia, un aumento de la ansiedad cuando los sujetos están realizando dicha tarea difícil.

Sin embargo, hay un parámetro, los tiempos de fonación, que parece contradecir lo encontrado con los anteriores. De acuerdo con Hicks (1979) y Hollien (1980), la ansiedad tiende a incrementar los tiempos medios de fonación. En ese supuesto, inicialmente la tarea de deletreo inverso induciría una mayor ansiedad, pues sus respuestas vocales tienen inicialmente mayores tiempos medios de fonación. Además, encontramos que en el curso de la realización de la repetición monovocálica los tiempos medios de fonación se incrementan fuertemente, en más del 50 %. Mientras que durante la realización de la tarea compleja, deletreo inverso, donde los niveles de ansiedad de los sujetos tienden a aumentar, los tiempos medios de fonación apenas varían, tal como se representa en la figura 1.

FIGURA 1



Evolución en el transcurso de la realización de las tareas de deletreo inverso y de repetición monovocálica de los tiempos medios de fonación (TMF) y de las intensidades promediadas de fonación (IMF).

Nosotros opinamos que los tiempos de fonación están muy relacionados con el contenido semántico del mensaje emitido y, por tanto, son muy sensibles al tipo de tarea. Hollien en su modelo de ansiedad se apoyó en los resultados encontrados por Hicks los cuales comparaban una conversación normal, situación de baja ansiedad, con una conversación en público, situación de alta ansiedad. Es presumible que las duraciones de las fonaciones, al igual que ocurre con el volumen de voz, fueran simplemente mayores por ser en público. De hecho, no pudieron replicar estos resultados cuando a los sujetos se les estudiaba en situaciones de laboratorio; en las cuales se inducía la ansiedad mediante descarga eléctrica. Nuestro caso también sería explicable por la diferencia en las tareas experimentales porque aunque las respuestas verbales son iguales, vocales /a/, proceden de tareas distintas que pueden imprimir duraciones de fonación diferentes sin, en un principio, ser debidas a distintos niveles de ansiedad. Sin embargo, este mismo argumento no puede aplicarse para el aumento que ocurre en los tiempos de fonación, durante la ejecución de la misma tarea de repetición monovocálica. Es más, sugiere la posición casi opuesta: que la ansiedad introduce titubeos y pequeñas pausas que tienen como consecuencia un acortamiento medio de las fonaciones.

De todas formas, no tenemos referencias bibliográficas para llegar a afirmar lo anterior con suficiente garantía científica. Pero sí tenemos otros resultados que también parecen apoyarlo. Estudiando la correlación entre los parámetros vocales y las puntuaciones en la escala STAI en su forma de ansiedad de rasgo, observamos que los tiempos medios de fonación correlacionan negativamente ($r = -0,559$, $p < 0,01$) con dichas puntuaciones. Es decir, los sujetos con personalidad más ansiosa producen fonaciones más cortas.

Siguiendo con las relaciones entre puntuaciones en el STAI y parámetros vocales, nos aparece un resultado bastante interesante para discutir, los sujetos de personalidad más ansiosa disminuyeron en menor medida su nivel de ansiedad durante la repetición monovocálica que los sujetos con personalidad menos ansiosa lo cual es puesto de manifiesto fundamentalmente por la correlación negativa con los incrementos del techo de la frecuencia fundamental ($r = -0,5416$, $p < 0,01$). Esto se puede interpretar como que los sujetos de personalidad ansiosa no alcanzan mayores niveles de ansiedad, al comienzo, que los sujetos no ansiosos sino que se relajan menos que los de personalidad no ansiosa durante la ejecución de las tareas sencillas. Es decir, tienen un menor control a largo plazo de la ansiedad. Este hecho sería explicado por autores como Morris y Engle (1981) o Kent y Jambunathan (1989) por sufrir dichos sujetos mayor número de intrusiones cognitivas negativas, así como, una pérdida de control sobre éstas.

Por otro lado, queremos hacer destacar que el análisis correlacional nos ha puesto de manifiesto otra aparente inconsistencia que no es fácil de explicar. La intensidad promediada de fonación correlacionó negativamente en todas las situaciones, y en alguna de ellas bastante significativamente ($r = -0,7337$, $p < 0,001$), con las puntuaciones del STAI en su forma de rasgo. Es decir, los sujetos con personalidad más ansiosa pronunciaban las vocales con una menor intensidad. Esto pudiera contradecir la hipótesis generalmente admitida de que la ansiedad incrementa la intensidad vocal, siempre en el supuesto que los sujetos de personalidad más ansiosa alcan-

zarían estados de ansiedad más altos durante la ejecución de las tareas, al menos en los intervalos finales. Esta hipótesis también pudiera estar corroborada por las variaciones que hemos encontrado en el curso de la realización de la tarea de deletreo inverso, tarea ansiógena, donde se produjeron incrementos significativos de la intensidad promediada de las fonaciones tal como se aprecia en la figura 1.

Nosotros opinamos que las correlaciones negativas entre la intensidad y las puntuaciones del STAI son las que explican adecuadamente la relación. Es decir, a mayor ansiedad corresponde menor intensidad vocal siempre que no influyan otras variables. Si dijimos anteriormente que los tiempos de fonación eran sensibles al contenido semántico de los mensajes, ahora añadimos que la intensidad vocal es especialmente sensible a la superposición de otras emociones. Por ejemplo, la ira tiende a subirla mientras que la tristeza o el miedo tiende a bajarla (Van Bezooeyen 1984, Scherer 1986). También depende de variables situacionales: hablar en público o en privado, así como, de componentes fisiológicos. En la figura 1 se pone de manifiesto un fenómeno sistemático en la fisiología del aparato fonatorio que aparece siempre en los ejercicios vocales continuos: el aumento paulatino de la energía vocal emitida, al menos durante los primeros minutos, independientemente de la tarea que el sujeto está realizando. Como podemos ver en la figura 1, para cualquiera de las dos tareas el producto intensidad por tiempo es mayor en el intervalo final que en el inicial. Si en la tarea generadora de ansiedad, deletreo inverso, los tiempos medios de fonación se mantuvieron o tendieron a disminuir por el incremento de ansiedad, para que la energía vocal aumentase, la intensidad vocal, necesariamente, tuvo que aumentar. Entonces podemos sugerir, como explicación de la inconsistencia con las correlaciones con el STAI, que el aumento de la intensidad vocal en la tarea de deletreo inverso se produjo no porque la ansiedad se incrementara si no porque la energía vocal aumentó.

CONCLUSIONES

Se ha puesto experimentalmente de manifiesto, en contra de nuestra hipótesis de partida, que la respuesta emocional de ansiedad inicialmente es similar, o no puede ser demostrada estadísticamente su diferencia, entre tareas cognitivas sencillas y complejas. Sin embargo, a lo largo de la ejecución, los sujetos sí son capaces de controlar su nivel de ansiedad y ajustarlo a las demandas de cada tarea. Así, en el deletreo inverso, tarea compleja, la ansiedad aumenta mientras que en la repetición monovocálica, tarea sencilla, la ansiedad se mantiene e incluso puede disminuir. Dado que la respuesta emocional de ansiedad está presente en una gran proporción de experimentos con sujetos humanos y que tiene una influencia tan considerable en el rendimiento y en muchas otras variables psicológicas, consideramos que debería ser controlada sistemáticamente pues nuestros resultados experimentales sugieren que puede variar significativamente en el transcurso de aquéllos y que los niveles de ansiedad finales dependen de las tareas que los sujetos estén realizando.

Comparando los sujetos que tienen una personalidad más ansiosa (según las puntuaciones en el cuestionario STAI en su forma de rasgo) con

los de personalidad menos ansiosa, hemos encontrado que la principal diferencia no estriba en que los primeros alcancen inicialmente mayores niveles de ansiedad de estado, si no que reducen en menor proporción o mantienen, respecto a los valores iniciales, sus niveles de ansiedad durante la ejecución de las tareas sencillas. Esto último podría explicar las discrepancias tradicionalmente encontradas entre las medidas psicofisiológicas y las puntuaciones en los cuestionarios de ansiedad.

Referencias

- ALMEIDA, A.; FLEISCHMANN, G.; HEIKE, G., y THORMANN, E. (1975). «Short time statistics of the fundamental tone in verbal utterances under psychic stress». *Eight International Congress of Phonetic Sciences*, Leeds, England.
- BARLAND, G. H. (1973). «Use of voice changes in the detection of deception». *Acoustical Society of America*, Los Angeles, Estados Unidos.
- BRENNER, M.; BRANSCOMB, H. H., y SCHWARTZ, G. E. (1979). «Psychological stress evaluator-two tests of a vocal measure». *Psychophysiology*, 16: 351-357.
- BRENNER, M.; SHIPP, T., y MORRISSEY, P. (1983). «Voice measures of psychological stress-laboratory and field data». *Iowa Conference on Voice*, Iowa, Estados Unidos.
- BROCKWAY, B.F. (1979). «Situational stress and temporal changes in self-report and vocal measurements». *Nursing Research*, 20-24.
- EYSENCK, M. W. (1983). «Anxiety and individual differences». Nueva York, NY: Stress and fatigue in human performance, en Hockey, G. R. J. (ed.): *Stress and Fatigue in Human Performance*, Nueva York, NY: John Wiley & Sons Ltd., 273-298.
- EYSENCK, H. J., y EYSENCK, M. W. (1987). «Personalidad y diferencias individuales», ed. Pirámide, S. A. Madrid, 286-306.
- FOA, E. B.; MCNALLY, R., y MURDOCK, T. B. (1989). «Anxious mood and memory». *Behav. Res. Ther.*, 27 (2): 141-147.
- FRIEDHOFF, A. J.; ALPERT, J. M., y KURTZBERG, R. L. (1964). «An electro-acoustical analysis of the effects of stress on voice». *J. Neuropsychiatry*, 5: 265-272.
- GOTTSCHALK, L. A., y FRANK, E. C. (1967). «Estimating the magnitude of anxiety from speech». *Behavioral Science*, 12: 289-295.
- HECKER, M. H. L.; STEVENS, K. N.; BISMARCK, G., y WILLIAMS, C. E. (1968). «Manifestations of task-induced stress in the acoustic speech signal». *The Journal of the Acoustical Society of America*, 44 (4): 993-1001.
- HICKS, J. W. (1979). «An acoustical/temporal analysis of emotional stress in speech». *Tesis doctoral no publicada*. University of Florida, Gainesville, FL.
- HOLLIER, H. (1980). «Vocal indicator of psychological stress», en Wright, F.; Bahn, C., y Reiber, R. W. (eds.): *Forensic Psychology and Psychiatry*, Nueva York: Annals New York Academy of Sciences, 47-72.
- HURLEY, P. A. (1983). «Communication variables and voice analysis of marital conflict stress». *Nursing Research*, 32: 164-169.
- INBAR, G. F.; EDEN, G., y KAPLAN, M. A. (1977). «Frequency modulation in the human voice and the source of its mediation». *Science and Engineering*, 213-219.
- JÜRGENS, V. (1983). «La voz como vehículo de expresión emocional y su regulación por el sistema nervioso central». *El Hexágono de Roche*, 6 (2): 7-11.
- KENT, G., y JAMBUNATHAN, P. (1989). «A longitudinal study of the intrusiveness of cognition in test anxiety». *Behav. Res. Ther.*, 27 (1): 43-50.
- KURODA, I.; FUJIWARA, O.; OKAMURA, N., y UTSUKI, N. (1976). «Method for determining pilot stress through analysis of voice communication». *Aviation, Space, and Environmental Medicine*, 47: 528-533.
- LALLJEE, M., y COOK, M. (1975). «Anxiety and ritualized speech». *British Journal of Psychology*, 66: 299-306.
- MENACHEM, R. (1983). «La voix et la communication des affects». *Année Psychologique*, 83: 537-560.
- MEULEMANS, G. (1974). «Mesure de la tension émotionnelle, d'après le comportement verbal du locuteur: étude théorique et statistique». *Bulletin du CERP*, 3: 141-151.
- MUSUMECI, M. A. (1975). «Speech disturbances as a function of stress induced anxiety in children». Tesis doctoral: Fordham University.
- RAGSDALE, J. D. (1976). «Relationships between hesitation phenomena, anxiety and self-control in a normal communication situation». *Language and Speech*, 14: 257-265.
- SCHERER, K. R. (1979). «Non linguistic vocal indicators of emotion and psychopathology», en Izard, C. E.: *Emotion and Psychopathology*. Nueva York: Plenum Press, 1979.

- SCHERER, K. R. (1986). «Voice, stress and emotion», en Appley, M. H., y Trumbull, R. (eds.): *Dynamic of stress: Physiological, psychological, and social perspectives* (págs. 157-179). Nueva York: Plenum Press.
- SHAPIRO, K. L., y LIM, A. (1989). «The impact of anxiety on visual attention to central and peripheral events». *Behav. Res. Ther.*, 27 (4): 345-351.
- SIMONOV, E. V., y FROLOV, M. V. (1973). «Utilization of human voice for estimation of man's emotional stress and state of attention». *Aerospace Medicine*, 44: 256-258.
- SMITH, G. A. (1977). «Voice analysis for the measurement of anxiety». *British Journal of Medical Psychology*, 50: 367-373.
- SPIELBERGER, C. D.; GORSUCH, R. L., y LUSHENE, R. E. (1970). «Manual for the State-Trait Anxiety Inventory». Palo Alto: Consulting Psychologist's Press.
- TOLKMITT, J., y SCHERER, K. R. (1986). *Effect of experimentally induced stress on vocal parameters*, Manuscrito no publicado.
- VAN BEZOOYEN, R. (1984). *Characteristics and recognizability of vocal expression of emotion*, Dordrecht: Foris publications.
- WEINER, H. (1989). «El papel del estrés en la salud y la enfermedad». 10.º Congreso Mundial del Colegio Internacional de Medicina Psicosomática, Madrid, España.
- WILLIAMS, C. E., y STEVENS, K. N. (1972). «Emotions and speech: Some acoustical correlates». *Journal of the Acoustical Society of America*, 52: 1238-1250.

APENDICE

Dado que usamos un conjunto de términos técnicos no muy extendidos dentro del área de Psicología Social, vamos a intentar dar una explicación muy sucinta de algunos de los más importantes.

Frecuencia fundamental: es el número de veces que abren y cierran las cuerdas vocales en cada segundo, cuando pronunciamos, por ejemplo, una vocal. Es el concepto técnico de lo que cualitativamente se denomina «tono de la voz». Esta frecuencia, que conocemos a través de la onda de presión del aire que alcanza al micrófono, suele estar variando. Si medimos muchos valores podemos calcular su media (F_0), las desviaciones típicas y diversos percentiles.

Tiempo de fonación: en este experimento en el que se manejan vocales aisladas, los tiempos de fonación reflejan, simplemente, la duración de dichas vocales. Son tiempos mucho más cortos de lo que pudieran imaginar (ver tabla II).

Intensidad de fonación: cuando pronunciamos una vocal, inicialmente el volumen de nuestra voz tiene una fuerte subida, posteriormente se mantiene y acaba decreciendo. El valor medio durante el tiempo de fonación es lo que llamamos intensidad promedio de fonación. En vez de mirar globalmente la intensidad, podemos analizar los valores instantáneos de pico que alcanza cada vez que se abren las cuerdas vocales. De estos valores máximos podemos extraer diversos estadísticos, así como, las variaciones entre valores correlativos y que llamamos variabilidad ciclo a ciclo de la intensidad máxima.

Distribución espectral: aparte de la onda fundamental generada por las cuerdas vocales, aparecen otras ondas debidas a las resonancias y antiresonancias del tracto vocal. La distribución espectral nos muestra la energía (tamaño) que tienen todas las ondas, cada una con su frecuencia, que componen la señal compleja de la voz. Hay unas ondas que poseen una energía especialmente grande y a éstas les llamamos formantes. Tienen una importancia vital en el lenguaje ya que sus frecuencias relativas son las que nos permiten distinguir una vocal de otra. En teoría se suelen analizar hasta cinco formantes pero las vocales quedan definidas suficientemente con los tres primeros.