

LA EXPRESIÓN DE LA EMOCIÓN A TRAVÉS DE LA CONDUCTA VOCAL

C. MUÑOZ ABAD; A. JIMÉNEZ FERNÁNDEZ
Universidad Autónoma de Madrid

Resumen

El estado emocional del hablante puede ser transmitido fiablemente a través de la conducta vocal; los cambios que se producen en dicho estado son detectados y cuantificados en el producto acústico final, ya sea a través de una determinada instrumentación o en función de las valoraciones realizadas por jueces. El objetivo que se plantea, por tanto, el presente artículo es doble. Por una parte se revisan de forma sistemática aquellas investigaciones interesadas en la detección de marcadores emocionales en la conducta vocal; se analizan dichos trabajos en función de un esquema básico común, destacando las diferencias metodológicas existentes entre los mismos, así como las ventajas e inconvenientes, resultantes de optar por un método determinado. Se constata así la existencia de una gran variación entre dichos estudios en función de la muestra de voz elegida, método de inducción de la emoción, codificadores, registro y análisis del producto vocal y presentación de los datos. No obstante, se detecta una convergencia de resultados sobre los parámetros vocales característicos de las emociones investigadas de forma sistemática (alegría, tristeza, miedo y cólera). En segundo lugar, se recogen los trabajos dedicados al estudio de los marcadores vocales de las dimensiones emocionales; tales investigaciones se analizan en función de los dos objetivos principales que se plantean: la identificación de las dimensiones que en función de un conjunto de claves vocales transmiten de forma diferencial la información afectiva, por un lado; y por otra parte, la identificación de tales marcadores y su relación con las dimensiones correspondientes.

Abstract

Speaker's emotional state can be transmitted by vocal behavior and it is possible to detect them in the acoustical final outputs through specific instrumentation or by judgment assessment techniques. Our goals in this work are two. We review the emotional markers in vocal behavior researches, in this way we analyzed the papers following the same basic system, focussing the methodological differences between them. There are a very high variability between studies in relation with the voice samples used, emotion elicited method, coders, vocal outputs recordings and analysis and data presentations. However, we detected different vocal acoustic parameters that identified the emotional states most investigated in the field (joy, sadness, fear and anger). On the other hand, we review dimensional works of the vocal markers. These researches had two goals: the identification of dimensions in terms of vocal cues and the identification of markers and their relation with these dimensions.

Introducción

La conducta vocal es sin duda la conducta comunicativa más importante. No sólo transmite mensajes lingüísticos, sino que empleando el mismo mecanis-

mo de producción transmite de forma simultánea mensajes de carácter no verbal, siendo el producto de dicha conducta físico (acústico) y por tanto objetivo (Del Pozo, Martínez, Jiménez, Zaplana y Delgado, 1981) en parámetros como la frecuencia fun-

damental, intensidad, energía, patrones temporales, etcétera. Tales parámetros actúan como marcadores de las características biológicas, psicológicas y sociales del emisor (Scherer y Giles, 1979), lo que no impide que cumplan además funciones lingüísticas específicas.

Por esta razón no es de extrañar que se haya considerado la conducta vocal como un indicador del estado emocional sujeto. La influencia del mismo sobre el proceso de producción vocal queda patente en los cambios que se producen en el patrón respiratorio y en la tensión de los músculos implicados en este proceso (Fry, 1976). Estos cambios son susceptibles de estudio, ya que afectan tanto a la tasa de vibración de las cuerdas vocales como a la forma del tracto vocal, determinando ésta, junto con el tipo de fonación, la cualidad de la voz (distribución de la energía en el espectro vocal). Este hecho permite la detección y cuantificación de dichos cambios emocionales en el producto acústico final, bien a través de una instrumentación adecuada o recurriendo a valoraciones realizadas por jueces.

Fiabilidad de los parámetros vocales y codificación emocional de la conducta vocal

Las investigaciones interesadas en el estudio de la conducta vocal han puesto de manifiesto que los marcadores emocionales del habla pueden ser valorados adecuadamente. Características vocales tales como el tono, la sonoridad o la tasa de articulación pueden ser evaluadas con elevada fiabilidad empleando sujetos entrenados, ligeramente entrenados o grupos de sujetos no entrenados (siempre y cuando se disponga de un grupo numeroso) (Constanzo, Markel y Constanzo, 1969; Davitz, 1964; Duncan y Rosenthal, 1968; Huttar, 1968; Markel, 1965; Scherer, 1978; Van Bezooeyen, 1984).

Dicha fiabilidad depende, no obstante, de determinados factores que es preciso considerar. Hay que tener en cuenta el contenido verbal, ya que si éste se mantiene constante resulta más fácil evaluar los parámetros considerados. Importa igualmente la longitud de las expresiones, al verse facilitada la tarea evaluativa para determinadas características vocales (sonoridad, rango y nivel del tono) cuando las expresiones son breves. Por último, señalar que la fiabilidad puede ser calculada a través de diversos procedimientos, que producen resultados diferenciales.

Los estudios de reconocimiento de las emociones han demostrado que éstas pueden ser identificadas a través de la voz, bien recitando las letras del abecedario (Davitz y Davitz, 1959; Dusenbury y Knower, 1938), leyendo números (Pfaff, 1954), susurrando las letras para eliminar el efecto del tono (Knower, 1941); o incluso aumentando sistemáticamente la proporción señal/ruido mientras el sujeto

lee distintas frases (Pollack, Rubenstein y Horowitz, 1960).

El principal inconveniente que presentan estos estudios es el de no permitir conocer qué marcadores contribuyen a la discriminación de las expresiones vocales emocionales. Dicho problema ha sido abordado de diferentes maneras. Soskin y Kauffman (1961), empleando un filtro de paso bajo determinaron la existencia de un conjunto de claves situadas en una banda de frecuencia entre 100 y 550 Hz, en base a las cuales los jueces realizaban inferencias similares sobre los estados emocionales codificados en las correspondientes muestras de voz.

Resultados similares fueron obtenidos por Kramer (1964) y Ross, Duffy y Cooker (1973) empleando esta misma técnica; incluso empleando simultáneamente diferentes técnicas como el filtrado y *random splicing* (empalme aleatorio de la cinta grabada), las emociones pueden ser identificadas en las distintas expresiones. No han sido muchas las investigaciones llevadas a cabo para estudiar la importancia diferencial de los parámetros vocales que permiten una discriminación entre expresiones vocales emocionales. En el trabajo de Lieberman y Michaels (1962) se estudió el efecto que sobre el reconocimiento emocional tenía la manipulación de dos parámetros acústicos, la frecuencia fundamental y la amplitud. Dos conclusiones se extraen del mismo: se constata, por un lado, la ausencia de un único correlato acústico de los estados emocionales, conclusión ratificada posteriormente por los resultados de las investigaciones de Jiménez (1987), Mallo (1986) y Scherer, Ladd y Silverman (1984). Por otra parte, se señala la existencia de una contribución diferencial de los parámetros acústicos en las expresiones emocionales; conclusión igualmente reforzada por el estudio antes mencionado (Scherer et al., 1984).

Se reconoce, por tanto, la existencia de un conjunto de claves vocales no verbales que, al actuar en forma conjunta, permitirían transmitir parte de la información sobre el estado emocional del hablante, con independencia de la información transmitida verbalmente. Sin embargo, es preciso distinguir entre distintos tipos de claves no verbales (Scherer et al., 1984), así la cualidad de la voz (relacionada con la distribución espectral de la energía) podría ser transmitida con independencia del tipo de contenido verbal, al igual que el nivel y rango de la frecuencia fundamental; mientras que el contorno de entonación dependería del tipo de frase o información lingüística a codificar.

Indicadores vocales de las emociones

Las investigaciones llevadas a cabo con el fin de detectar la existencia de marcadores emocionales en la conducta vocal, difieren metodológicamente en diversos aspectos, presentando, no obstante, un esquema básico (Fig. 1) común a la mayoría de ellas. Según este esquema, uno o varios codificado-

res expresan un conjunto de emociones verbalmente, cuyo contenido lingüístico ha sido previamente especificado. La muestra de voz es registrada y analizada, bien acústicamente, empleando procedimientos instrumentales para extraer los parámetros acústicos que supuestamente actúan como marcadores emocionales; bien de forma perceptual, siendo uno o varios jueces quienes valoran dichos parámetros.

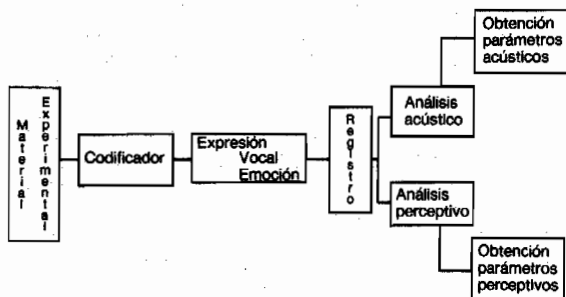


Figura 1. Modelo de investigación de los marcadores emocionales en la conducta vocal seguido por una gran parte de los trabajos empíricos del campo.

Este esquema nos sirvió como punto de partida para realizar la revisión de trabajos, dado que en él se destacan todos aquellos aspectos en los que difieren los estudios realizados, y que actúan como fuente de variación en los resultados, dificultando en algunos casos la comparación entre los mismos.

Muestra de voz

Una de las principales cuestiones que debe plantearse el investigador interesado en el estudio de los indicadores vocales emocionales es cómo controlar el contenido verbal. Por tanto, el primer elemento a considerar es la muestra de voz elegida. Parecería lógico, a primera vista, suponer que el método más adecuado para estudiar las expresiones emocionales sería emplear expresiones registradas en la vida real y producidas espontáneamente. Sin embargo, muy pocos estudios (Huttar, 1968; Williams y Stevens, 1972, 1981) han elegido *muestras de habla espontánea*. En otros casos, la muestra de voz no ha sido obtenida en situaciones reales, pero contienen, no obstante, elementos naturales (Markel, Bein y Phillis, 1973).

La selección de muestras de habla registradas en este tipo de situaciones implica no pocos inconvenientes que han de ser tenidos en cuenta. En estas muestras el contenido verbal no está sometido a ningún tipo de control, por lo que resulta difícil precisar en qué medida una emoción es transmitida únicamente a través de las claves no verbales del habla. Plantean asimismo dificultades de obtención, dado que, y en palabras de Scherer, «las emociones

son experiencias privadas, que en la mayoría de las culturas, al menos en sus manifestaciones extremas, han de ser suprimidas en público» (1981, p. 201).

La necesidad de registros de alta calidad, especialmente en aquellos casos en que van a realizarse análisis acústicos precisos, es otra de las desventajas a tener en cuenta cuando se emplean expresiones obtenidas en la vida real.

Otro problema que se plantea es la dificultad de saber con exactitud qué emoción concreta está experimentando el hablante. Esta cuestión es especialmente relevante en las investigaciones transculturales, puesto que los estímulos que elicitaban una emoción concreta pueden variar de una cultura a otra.

De igual modo, resulta difícil obtener emociones puras en las situaciones cotidianas, ya que suelen presentarse en cada situación varias emociones simultáneamente.

Un último inconveniente es que está relacionado con la réplica de los resultados. Las expresiones emocionales espontáneas son casos únicos, tanto en términos del hablante como de la situación; es difícil obtener la misma expresión emocional para el mismo hablante en otra situación o expresiones de la misma emoción para distintos sujetos.

A la vista de todos los problemas señalados anteriormente, parece lógico que apenas se hayan empleado registros de situaciones naturales, recurriéndose a otros métodos de obtención de las muestras de voz. Así, diversos investigadores han optado por escoger muestras cuyo *contenido verbal carecía de significado*, empleando palabras (Bortz, 1966; Skinner, 1935) o letras del alfabeto (Green y Cliff, 1975; Kaiser, 1962).

Este tipo de material plantea igualmente determinados inconvenientes que creemos es preciso considerar. Por un lado se cuestiona su naturalidad, dado que, al ser una muestra de voz sin significado, es probable que pueda provocar una exageración estereotípica en el codificador. Por otra parte, al ser muestras de habla excesivamente cortas pueden no contener todas las posibles claves relevantes para la comunicación emocional.

En otros casos, los investigadores han planteado el control del *contenido verbal*, manteniéndolo *constante*, esperando así que la emoción sea transmitida únicamente a través del canal no verbal. Los estudios realizados han sido de tres tipos en función de dicho contenido.

En el primero de los casos considerado se seleccionaban una o más frases, cuyo contenido se supone neutro, siendo emitidas simulando diferentes estados emocionales (Jiménez, 1987; Van Bezoozen, 1984).

En otros trabajos, la expresión elegida estaba enmarcada en distintos contextos (o similares), según el tipo de emoción, con el objetivo de hacer la tarea del codificador más natural. Tales son las muestras de voz seleccionadas en los trabajos de Constanzo et al. (1969), Davitz (1964), Fairbanks y Hoaglin (1941), Fairbanks y Pronovost (1939).

En el estudio de Williams y Stevens (1972, 1981) se siguió un procedimiento más complejo, elaborándose un pequeño diálogo para ser representado por tres actores en un escenario. Las situaciones emocionales estaban relacionadas para permitir cambios lógicos de unas a otras. En el diálogo se repetían una serie de frases (frases control) sobre las que se realizaba el análisis acústico; extrayéndose medidas de la frecuencia fundamental, intensidad, tasa de habla (medida en sílabas por segundo), así como un análisis visual del espectrograma y una descripción del contorno de entonación.

Igualmente se ha empleado como material estándar no una frase, sino un texto. La ventaja de este tipo de muestras de voz reside en su extensión, ya que, al ser suficientemente largas, permiten la extracción de todos los posibles parámetros relevantes para la comunicación emocional. Sin embargo, presentan, a nuestro juicio, un grave inconveniente, ya que, al ser expresiones muy largas, difícilmente podrá el codificador mantener la misma emoción a lo largo de todo el texto. Ejemplos de tales estudios son los realizados por Feldstein, Jaffe y Cassotta (1968) y Scherer, London y Wolf (1973).

Un trabajo especialmente interesante es el realizado por Levin y Lord (1975), dado que es uno de los primeros estudios junto con el de Scherer et al. (1973), en el que se emplean técnicas de procesamiento digital de la señal. La muestra de voz, en este caso, eran palabras elegidas libremente por los propios codificadores, quienes elegían igualmente la emoción a expresar.

Un estudio diferente respecto a la muestra de voz es el de Scherer (1974), en el que se utilizaban *estímulos artificiales* generados por un sintetizador, variando sistemáticamente cinco parámetros. Cada emoción aparece caracterizada por un conjunto de parámetros con determinados valores; definiendo igualmente las interacciones entre ellos algunas de las emociones.

Técnicas de enmascaramiento del contenido

Otra de las alternativas que se ha planteado para separar el canal verbal del vocal ha sido la de emplear técnicas de enmascaramiento del contenido.

La técnica más simple consiste en poner el *registro hacia atrás*. El contenido resultante es ininteligible, alterándose el contorno de entonación original; sin embargo, se crea un nuevo contorno con un énfasis muy elevado, generalmente al final de la frase (Scherer et al., 1984). Otra técnica de enmascaramiento es la desarrollada por Scherer (1981), denominada *random splicing*. Básicamente esta técnica consiste en eliminar las pausas de una muestra de voz, segmentar el resto del registro y unir aleatoriamente los fragmentos. Si éstos son suficientemente breves (entre 150 y 300 ms), el flujo de voz resultante es ininteligible, quedando preservada la información respecto al tono (nivel y rango), amplitud (nivel

y rango) y tasa de producción. Mientras que otras claves relevantes para la comunicación del significado emocional, tales como rasgos temporales y patrones de entonación, quedan eliminadas.

Sin embargo, presenta ciertos inconvenientes, ya que parece crear un nuevo contorno de entonación; e incluso puede dejar intactos fragmentos de palabras y crear la ilusión de otras nuevas.

Una tercera técnica consiste en pasar el habla registrada por un *filtro de paso bajo*, que elimina las altas frecuencias, portadoras de gran parte de la información verbal y fundamentales para el reconocimiento del habla, permaneciendo los sonidos de baja frecuencia, portadores de la información emocional (Soskin y Kauffman, 1961). El punto de corte elegido oscila alrededor de los 500 Hz. El inconveniente, sin embargo, reside en que las altas frecuencias también parecen contribuir a la percepción de la cualidad de la voz, por lo que se perdería información relevante emocionalmente. Además, si la misma muestra de habla se presenta repetidamente, los sujetos pueden llegar a entender parte del contenido verbal (Kramer, 1964). Estudios en los que se han empleado filtros de paso bajo son los de Kramer (1964), McCluskey, Niemi y Albas (1978), Scherer et al. (1972) y Soskin y Kauffman (1961).

Métodos

Tal como señalábamos al principio, según el esquema propuesto, los codificadores han de expresar vocalmente un conjunto de emociones, habiéndose empleado, según los trabajos, diversos métodos para provocar la emoción deseada.

En aquellos estudios en los que la muestra de voz había sido recogida en situaciones reales (Huttar, 1968; Williams y Stevens, 1972), las *emociones surgen espontáneamente*. A los inconvenientes mencionados al tratar la utilización de muestras de habla recogidas en situaciones naturales, hay que añadir el problema de la falta de control sobre los factores que determinan la aparición de la conducta emocional. Éste es un aspecto esencial a tener en cuenta si se quiere llevar a cabo un estudio sistemático de la correspondencia entre los valores de los parámetros acústicos y los diferentes estados emocionales.

La mayoría de los investigadores, tal como ha sucedido en el área de la expresión facial, han optado por emplear *expresiones simuladas* en situaciones de laboratorio (Zuckerman, Hall, De Frank y Rosenthal, 1976), debido a los múltiples problemas que plantea el registro de emociones espontáneas.

Esto permite controlar las condiciones de registro, utilizando equipos de alta calidad (como en los trabajos de Jiménez, 1987; Levin y Lord, 1975; Scherer et al., 1973; Van Bezooeyen, 1984). Es igualmente posible establecer una adecuada comparación entre las descripciones emocionales, siendo además menos probable obtener emociones mezcladas.

Williams y Stevens (1972) compararon el registro original del locutor de radio que describía el desastre del Hindenburg, con el de un actor al que se le pedía leer una versión escrita del comentario del locutor simulando su papel. El análisis acústico mostró que tanto en el habla original como durante la simulación aparecían variaciones irregulares de la frecuencia fundamental, un incremento de la mediana de la frecuencia fundamental y un rango mayor de la misma (comparando los registros de antes y después del golpe).

Los fenómenos acústicos descritos eran, en los tres casos, más claramente visibles en el registro simulado que en el espontáneo; lo que puede ser considerado como un índice de la existencia de características comunes entre las expresiones simuladas y las espontáneas.

Este tipo de situaciones simuladas permite además al experimentador provocar la conducta deseada repetidamente para replicar los resultados, así como manipular aspectos específicos de la situación experimental. Ha sido, por tanto, uno de los métodos más empleados en los distintos estudios.

Otros métodos como la *inducción* han sido empleados más esporádicamente. Este método consiste en provocar el estado emocional deseado en el sujeto, haciéndole leer un poema o escuchar un fragmento musical, antes de la emisión de la expresión emocional. No ha sido muy utilizado (Sedláček y Sychra, 1963; Skinner, 1935; Zuberbier, 1957), ya que el rango de emociones que puede ser inducido es limitado (en el trabajo de Skinner sólo se estudiaban alegría y tristeza; y en el de Sedláček y Sychra, alegría, tristeza y miedo); añadiéndose el inconveniente de que los efectos de la inducción pueden no ser lo suficientemente fuertes como para manifestarse en la expresión vocal.

Codificadores

Con respecto a los codificadores, es interesante considerar las diferencias respecto al número de ellos empleado en los diversos trabajos, así como el sexo, experiencia y entrenamiento de los mismos.

El número de sujetos empleados como codificadores de la expresión emocional presenta una gran variabilidad. Desde los estudios de Eldred y Price (1958); Green y Cliff (1975); Huttar (1968); Jiménez (1987); Scherer et al. (1973); en los que un solo sujeto expresa todas las emociones; hasta otros como los de Feldstein et al. (1968); Markel et al. (1973); Sedláček y Sychra (1963); Zuberbier (1957), cuyo número de codificadores oscila entre veinte y cincuenta.

Aunque en algunos estudios el número de sujetos es muy elevado, apenas ha sido considerada la variabilidad entre los codificadores. Es necesario tener en cuenta tal aspecto, ya que, según se desprende del trabajo de Lieberman y Michaels (1962), los hablantes tienden a favorecer el uso de diferentes pa-

rámetros acústicos para transmitir la misma información emocional. Una adecuada consideración de tal variabilidad nos permitiría establecer semejanzas y diferencias entre los sujetos, lo que redundaría en beneficio de la generalización de nuestros resultados.

Sobre la *experiencia de los codificadores* hay que señalar que, aproximadamente en la mitad de los estudios revisados, se seleccionan como codificadores, actores profesionales o estudiantes de arte dramático, mientras que en el resto los sujetos carecen de experiencia en este terreno.

La selección de actores o estudiantes de arte dramático presenta una gran ventaja, ya que, dada su formación y entrenamiento, les será más fácil representar las expresiones emocionales requeridas, mostrando menor embarazo ante tal situación. Sin embargo, aunque la representación de los actores trate de enmarcarse en situaciones lo más naturales posibles (Scherer et al., 1972; Williams y Stevens, 1972), siempre existirán diferencias respecto a las expresiones espontáneas. El actor puede tender, incluso, a exagerar la emoción creando efectos acústicos que no aparecerían en expresiones vocales obtenidas en situaciones reales.

Una posibilidad contraria a considerar, señalada por Williams y Stevens (1981), es la de que aunque el actor pueda representar una emoción de tal forma que sea decodificada correctamente, pueden no producirse en él los cambios fisiológicos que acompañan a esta emoción en situaciones reales. Estos cambios podrían provocar determinados efectos acústicos en la voz, que no aparecerían en las expresiones vocales simuladas.

Tal posibilidad puede ser controlada empleando registros de determinadas respuestas fisiológicas (como, por ejemplo, la tasa cardíaca en el estudio de Simonov, Frolov y Taubkin, 1975; o la respuesta electrodérmica en el trabajo de Skinner, 1935), con el fin de obtener un criterio de la emocionalidad del sujeto.

Las diferencias entre codificadores en función del sexo apenas han sido tenidas en cuenta. En el trabajo de Davitz (1964) se obtuvo que la sonoridad, nivel del *pitch*, timbre y tasa de habla variaban de la misma forma para ambos sexos. Resultados similares se obtuvieron en el estudio de Constanzo et al. (1969) para los parámetros del *pitch*, duración y sonoridad; igualmente en los análisis realizados por Van Bezoooyen (1984) sólo se hallaron diferencias respecto a la laxitud laríngea y dureza de la voz: las expresiones vocales emocionales de los hombres eran percibidas más ásperas y con menor tensión laríngea que las de las mujeres, lo cual parece ser una diferencia característica entre la voz masculina y la femenina.

Sin embargo, en contraste con los resultados anteriores, Feldstein et al. (1968) obtuvieron diferencias significativas en función del sexo del codificador, en todos los parámetros estudiados (número de interrupciones en el flujo de voz, tasa de respiración, variabilidad del nivel de intensidad y duración de las pausas).

Registro y análisis

De acuerdo con el esquema del que partimos en esta revisión, una vez que la emoción ha sido transmitida vocalmente, es registrada y analizada, siendo éste otro de los aspectos tratado de forma diferencial por el investigador. La muestra de voz puede ser analizada y descrita *perceptualmente*, siendo los sujetos quienes valoran los parámetros (*pitch*, sonoridad, duración, timbre), por medio de un conjunto de escalas cuyo número de categorías varía en función del trabajo considerado.

El número de jueces elegido, su experiencia o entrenamiento previo difieren considerablemente de unos trabajos a otros. La fiabilidad de las descripciones perceptuales realizadas por los sujetos está modulada por determinados factores, ya mencionados al comienzo de este trabajo.

La descripción de las expresiones vocales puede ser realizada *acústicamente*, empleando variados sistemas de registro y análisis, desde los más sencillos como los oscilógrafos (Fairbanks y Hoaglin, 1939; Fairbanks y Pronovost, 1941; Skinner, 1935), que requieren mucho tiempo, ya que sólo pueden ser examinadas cada vez breves porciones de la onda acústica; hasta las técnicas más sofisticadas del análisis digital (Jiménez, 1987; Levin y Lord, 1975; Scherer et al., 1973), pasando por espectrógrafos, que permiten una representación tridimensional frecuencia-tiempo-amplitud; o registradores del nivel de la frecuencia fundamental y del nivel de la amplitud.

Los parámetros acústicos utilizados con más frecuencia para caracterizar las emociones han sido el nivel de frecuencia fundamental, la intensidad, duración, rango y variabilidad de la frecuencia fundamental y duración de las pausas. Tales parámetros han sido considerados de forma univariada; sin embargo, los resultados obtenidos por Jiménez (1987) sugieren que la codificación vocal de la emoción es un fenómeno fundamentalmente multivariado, siendo diseñada posteriormente una investigación (Muñoz, 1986) con el fin de contrastar dicha hipótesis.

Por último, otros trabajos han combinado ambos procedimientos, instrumental y perceptual, realizando un tipo de descripción para cada parámetro (Feldstein et al., 1968); o correlacionando los resultados obtenidos al emplear ambos procedimientos conjuntamente (Van Bezoooyen, 1984).

Presentación de los datos

Directamente relacionado con lo anterior, hay que considerar la presentación de los datos. Éstos se han presentado bien verbalmente, en términos de alto, bajo, medio (respecto al tono, por ejemplo) (Davitz, 1964; Markel et al., 1973); numéricamente, en Hz, semitonos, segundos, según los parámetros concretos (Fairbanks y Pronovost, 1939; Feldstein et al., 1968; Jiménez, 1987; Levin y Lord, 1975; Scherer et al., 1973; Williams y Stevens, 1972,

1981) o según posiciones en una escala (Scherer, 1974). No existe, pues, un criterio formal para interpretar los resultados, adoptándose en muchos casos un criterio arbitrario (Van Bezoooyen, 1984).

Es además importante tener en cuenta que en muchos estudios se comparan unas expresiones emocionales frente a otras, mientras que en otros se hace respecto a una expresión que supuestamente no expresa ningún tipo de emoción (Bortz, 1966; Davitz, 1964; Levin y Lord, 1975; Williams y Stevens, 1972). Este aspecto no queda claramente explicitado en todos los trabajos, por lo que puede ocurrir que lo que se considere como *pitch* alto (por ejemplo) en un estudio puede ser bajo en otro, dependiendo del punto de referencia escogido, así como del conjunto específico de emociones estudiado.

Por último, señalar que la mayoría de los datos se han obtenido, fundamentalmente, sobre expresiones emocionales producidas por hablantes estadounidenses y más esporádicamente sobre muestras europeas (alemanes, ingleses, franceses, españoles). Teniendo en cuenta, además, que sólo unas pocas emociones han sido estudiadas sistemáticamente, parece claro que es difícil generalizar los resultados, así como realizar comparaciones transculturales de los marcadores empleados en la codificación emocional.

Tampoco se han incluido en un mismo estudio codificadores de distintas lenguas, de cara al estudio de las diferencias entre los marcadores emocionales. Una excepción es el estudio de Fónagy y Magdics (1963), en el que se comparaban los valores del nivel y rango del *pitch* para expresiones de sorpresa, alegría, miedo, desprecio e ira, emitidas por codificadores de distintas lenguas (húngaro, alemán, inglés y francés).

A pesar de las diferencias existentes entre los distintos estudios respecto al diseño, metodología y técnicas de análisis, existe una sorprendente convergencia de resultados respecto a los parámetros vocales característicos de las emociones investigadas más sistemáticamente, que son ira, tristeza, miedo y alegría, y fundamentalmente las dos primeras. Tales resultados pueden ser comparados con las predicciones hipotéticas que desde el modelo de procesos componentes (MPC) de Scherer (1984) se plantean con respecto a los cambios que los diferentes estados emocionales provocarían en la conducta vocal (caracterizada acústicamente). Tales cambios dependen tanto de la secuencia valorativa realizada sobre el estímulo como de las variaciones que se producen en los restantes subsistemas.

La emoción de ira se caracteriza por un alto nivel de la frecuencia fundamental en casi todos los estudios (excepto en los de Fónagy y Magdics, 1963, y Jiménez, 1987, cuyo valor no se altera), extenso rango y gran variabilidad de la frecuencia fundamental, voz intensa y breve duración (excepto en el trabajo de Jiménez (1987), en el que presenta una mayor duración. Estos resultados se corresponden con las predicciones que desde el MPC se establecen para esta emoción, con excepción de los valores

para la frecuencia fundamental (F_0). Los datos empíricos sustentan un aumento de la F_0 media, mientras que desde el MPC se predice una disminución de la misma; para Scherer tal discrepancia podría deberse a que los sujetos empleados en los diferentes estudios para simular las expresiones emocionales produjeran formas extremas de cólera con una elevada activación y un bajo componente aséptico.

Tristeza presenta características opuestas: bajo nivel de la frecuencia fundamental (mientras que en el estudio de Jiménez, 1987, su valor no varía; mostrando un aumento en el trabajo de Van Bezoooyen, 1984); rango estrecho y pequeña variabilidad de la frecuencia fundamental, voz poco intensa y gran duración. Van Bezoooyen (1984), sin embargo, obtiene un aumento en el nivel del tono para el que no encuentra explicación. El patrón predicho desde el MPC para esta emoción se corresponde totalmente con los resultados empíricos obtenidos.

Con respecto a la emoción de miedo, se obtiene que el nivel de la frecuencia fundamental es alto (con la excepción de los trabajos de Fónagy y Magdics, 1963; y de Jiménez, 1987), extenso rango (el cual no se altera según los resultados del trabajo de Van Bezoooyen, 1984) y gran variabilidad de la frecuencia fundamental; la duración es breve (excepto en el estudio de Jiménez, 1987, en el que no varía) y la intensidad no varía (según los resultados de Van Bezoooyen, 1984, y Jiménez, 1987). Los resultados empíricos confirman las predicciones para los parámetros de la F_0 , pero no con respecto a los parámetros relacionados con la intensidad ni la duración (se predice un elevado valor para la intensidad y gran duración).

Alegria se caracteriza por un alto nivel de la frecuencia fundamental, gran variabilidad y rango extenso de la frecuencia fundamental, intensidad elevada y duración breve (mientras que ésta aumenta para Jiménez, 1987). Las predicciones teóricas del MPC para los parámetros de la intensidad y F_0 (baja frecuencia fundamental, rango estrecho, pequeña variabilidad de la F_0 , así como de la intensidad) difieren de los resultados empíricos; sin embargo, hay que señalar que la predicción de Scherer se establece para un estado emocional caracterizado por el encuentro con un estímulo placentero intrínsecamente, y/o la satisfacción de una necesidad o el logro de una meta. El sistema fisiológico tendería a reducir la activación con un predominio del sistema nervioso parasimpático, acorde con el sistema ejecutivo sobre consumación o descanso, lo que origina un estado en el receptor de placer, felicidad o alegría tranquila. Sin embargo, en la mayoría de los estudios el estado emocional considerado posee un alto componente activador, lo que originará cambios en la F_0 y en la energía con valores inversos al estado de felicidad descrito por Scherer.

Con respecto a las demás emociones estudiadas, las cuales cubren un amplio rango, no podemos extraer conclusiones definitivas, dado que cada una no se considera más que en uno o dos trabajos, y por tanto no pueden establecerse comparaciones adecuadas.

Indicadores vocales de las dimensiones emocionales

En una concepción dimensional de la emoción se asume que las mismas están caracterizadas por transiciones graduales; a diferencia de una concepción categórica, en la que las emociones son consideradas como entidades discretas claramente diferenciadas unas de otras. Asimismo se señala que las diferencias y similitudes entre emociones se describen en términos de distancias a lo largo de una serie de dimensiones. Éstas se corresponden casi siempre con las dimensiones principales del significado halladas por Osgood, Suci y Tannenbaum (1976), evaluación, actividad y potencia. En este enfoque, por tanto, se evalúan las emociones y sus correspondientes marcadores vocales, sobre escalas bipolares utilizadas en la técnica del diferencial semántico. Los trabajos que han investigado en esta línea se han planteado dos objetivos. Por un lado se ha pretendido la identificación de aquellas dimensiones que juegan un papel diferencial en la transmisión de la información afectiva, en función de un conjunto determinado de claves vocales; y por otra parte, se ha buscado la identificación de tales marcadores relacionándolos con las correspondientes dimensiones.

Del primero de los objetivos dan cuenta los trabajos de Apple y Hetch (1982) y Green y Cliff (1975), en los que fueron aisladas dos dimensiones, una evaluativa y otra de actividad. Igualmente, Mehriban (1972) detectó la existencia de una dimensión etiquetada como «responsividad vocal», que sería similar a la dimensión de actividad. Tales resultados han sido replicados posteriormente por Van Bezoooyen (1984), al aislar esta última dimensión.

El segundo de los objetivos, establecer la relación entre las dimensiones emocionales y los parámetros vocales, ha quedado recogido en diversos trabajos. Así, en el estudio de Scherer et al. (1972) se ha confirmado la existencia de un conjunto mínimo de claves vocales (nivel y rango de la frecuencia fundamental, rango de la intensidad y duración), que permiten transmitir las dimensiones del significado emocional.

En función de tales resultados se realizaron posteriormente dos estudios (Scherer, 1974; Scherer y Oshinsky, 1977), el primero de ellos de carácter piloto, en los que por medio de un sintetizador se variaban sistemáticamente diferentes parámetros de una secuencia de sonidos producidos artificialmente que eran valorados sobre escalas del diferencial semántico.

Las emociones agradables se caracterizaban por una duración breve, bajo número de armónicos, gran variación en el *pitch*, contorno descendente y pequeñas variaciones de la amplitud.

Las secuencias de tonos breves, con el nivel del *pitch* alto, gran número de armónicos, gran variación en el *pitch* y pequeñas variaciones en la amplitud, caracterizaban en general las emociones activas.

Mientras que para potencia los parámetros relevantes eran un gran número de armónicos, breve duración, nivel del *pitch* alto y contorno ascendente. Una combinación entre una gran variación de la amplitud y nivel del *pitch* alto o entre nivel alto del *pitch* y gran número de armónicos producía el mismo tipo de valoraciones. En general, parecía ser la duración el parámetro de mayor relevancia para la evaluación de los estímulos sobre la dimensión de actividad, especialmente destacado para las emociones con posiciones extremas en esta dimensión (aburrimiento, tristeza y alegría). Incluso ésta era la dimensión evaluada más fiablemente por los sujetos, lo cual estaba de acuerdo con los resultados de Davitz (1964), quien obtuvo una relación más consistente entre las claves de *pitch*, sonoridad, tasa y timbre; y el nivel de actividad evaluado subjetivamente, que respecto a cualquiera de las otras dos dimensiones.

El nivel de filtración (número de armónicos) resultó ser el segundo predictor por orden de importancia, relevante para las atribuciones de potencia. Mientras que la duración y el nivel de filtración contribuían en igual medida para marcar la dimensión evaluativa. Las secuencias con menor número de armónicos eran valoradas como más agradables y alegres; mientras que las que tenían un menor número de armónicos eran juzgadas más activas y potentes.

En otros estudios, los autores han empleado, en vez de sonidos producidos artificialmente, codificadores que expresaban distintas emociones mediante diversos tipos de contenido verbal. Sobre estas expresiones se evaluaba por un lado determinados parámetros acústicos, y por otro se juzgaba el significado emocional de las mismas por medio de las escalas del diferencial semántico, correlacionándose posteriormente ambas valoraciones.

En el trabajo de Bortz (1966), mediante un análisis de correlación canónica, se obtuvo que la intensidad era el predictor más poderoso; y el factor ritmo, el segundo más importante. Ambos factores correlacionaban con las dimensiones de actividad y evaluación.

Davitz y Davitz (1959) hallaron que entre la dimensión de actividad y las características del habla, *pitch*, sonoridad, timbre y tasa, existía una relación lineal positiva. Los sentimientos valorados subjetivamente por los jueces como activos (miedo o alegría) eran expresados con un *pitch* alto, voz sonora, tasa de habla rápida y timbre no resonante. Mientras que los evaluados subjetivamente como pasivos presentaban características opuestas. Huttar (1968) corroboró estos resultados, obteniendo una elevada correlación entre las variables acústicas de frecuencia fundamental, amplitud y duración, y la valoración del nivel de actividad de las emociones estudiadas.

Las dimensiones de potencia y evaluación parecían mantener una relación curvilínea con los parámetros mencionados, relación que en función de los datos obtenidos no pudo ser probada.

Las variaciones en el *pitch*, sonoridad y duración pueden ser evaluadas en base a segmentos como palabras o frases, pero para la identificación de di-

mensiones emocionales también es necesario tener en cuenta las relaciones entre segmentos, que permiten la valoración de patrones más complejos, como el ritmo y la inflexión.

Davitz (1964) señaló que estas claves o patrones caracterizarían la dimensión evaluativa. Las emociones agradables estaban acompañadas por un ritmo e inflexiones regulares, mientras que las emociones desagradables (como ira o tristeza) se caracterizarían por un patrón opuesto. Ritmo e inflexión permitirían además distinguir emociones similares en cualquiera de las otras características vocales.

Podríamos, por tanto, clasificar los rasgos vocales en función de las dimensiones emocionales: el nivel de actividad estaría codificado a través del *pitch*, tasa de habla y resonancia (la resonancia hace referencia a la apertura de las cuerdas vocales; una voz será resonante cuando las cuerdas vocales estén muy abiertas y relajadas). Mientras que el ritmo y la inflexión estarían más relacionadas con la dimensión evaluativa. Según Collier (1985), combinando estos rasgos se obtendrían cuatro patrones diferentes, que quedarían recogidos en la figura 2.

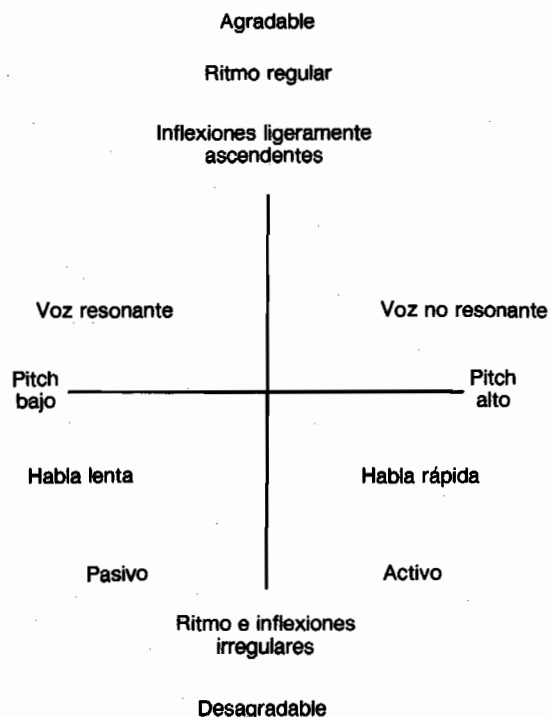


Figura 2. Clasificación bidimensional de la expresión emocional donde la medida de frecuencia reflejaría la dimensión de activación y el ritmo reflejaría la dimensión evaluativa conformando diversos tipos acústicos de conducta vocal.

Se han buscado diferentes explicaciones para la presencia de una correlación más prominente entre la dimensión de actividad y determinados rasgos vocales.

Abercrombie (1968) señaló que el habla es fundamentalmente una actividad muscular, y que por tanto estará sujeta a los mismos tipos de influencias que cualquier otra actividad muscular. La activación provoca un incremento en la tensión muscular (ya señalado por Lindsley, 1969, como correlato de determinadas emociones). Por un lado, existe un incremento en la tensión de los músculos respiratorios, y de la tasa de respiración, lo que produce un aumento de la intensidad vocal y un habla más rápida (Huttar, 1968). La tensión de los músculos respiratorios provoca un aumento de la presión subglotal, lo que determinará un mayor tensión de las cuerdas vocales, y por tanto una mayor tasa de vibraciones de éstas. Asimismo, puesto que la resonancia se basa en la apertura de las cuerdas vocales, y éstas durante las emociones activas están más tensas y cerradas, la voz resultante será menos resonante. Por último, la activación también altera la coordinación del habla; si es muy elevada, la voz pierde su ritmo y se hace más irregular.

Estas respuestas fisiológicas conducirían a una posible correlación entre la dimensión de actividad de la emoción y los parámetros vocales de incremento en el *pitch*, en la variabilidad del mismo, aumento de la intensidad y tasa de habla más rápida. Sin embargo, no debemos asumir, simplistamente, una relación unívoca entre actividad fisiológica y la actividad evaluada subjetivamente, o entre activación y determinadas expresiones vocales emocionales. Es necesario tener en cuenta, al menos en la conducta humana, el papel supresor e inhibitor de las normas sociales sobre las conductas emocionales, y su efecto modelador sobre las mismas.

Estos mismos cambios de los rasgos vocales, durante las emociones activas, han sido observados en diversas especies. Estudios evolutivos (Dimitrovsky, 1964; Van Bezoooyen, 1984) y transculturales (Albas, McCluskey y Albas, 1976; Van Bezoooyen, 1984) señalan igualmente que los errores en la identificación de las expresiones vocales emocionales parecen ser una función de la similitud en el nivel de actividad de tales emociones. Por tanto, la presencia de patrones vocales similares en niños, en diferentes culturas y otras especies distintas del hombre, podrían sugerir una estrecha relación entre las claves vocales que codifican la dimensión de actividad y los cambios fisiológicos provocados por la activación. Sin embargo, se precisan investigaciones más exhaustivas y sistemáticas, teniendo en cuenta que la expresión vocal emocional es una parte de todo un proceso comunicativo en el que los canales facial, postural y verbal juegan también un papel importante; e incluso no debemos olvidar que nuestros conocimientos fisiológicos y anatómicos son todavía incompletos en muchos aspectos.

Al igual que se predicen desde el modelo de procesos componentes de Scherer (1984) determinados cambios en la conducta vocal en función de los diferentes estados emocionales, se contemplan también transiciones entre los estados emocionales prototípicos. Se supone que tales transiciones estarán caracterizadas por cambios graduales en los

parámetros relacionados con los resultados de las exploraciones secuenciales específicas realizadas.

Por un lado, puede producirse una variación continua desde un registro bajo o de pecho (caracterizado por una baja frecuencia fundamental e intensidad elevada), el cual sería un indicativo de la conciencia de la propia fuerza, hasta un registro de cabeza o entonación en falsete, que señalaría la falta de confianza y poder del emisor. Esta transición dependerá de los resultados valorativos de la capacidad del organismo para enfrentarse con el estímulo. Fisiológicamente, el responsable del control de la frecuencia fundamental en el registro de cabeza sería el músculo cricotiroides, innervado por el nervio laríngeo externo; mientras que el músculo vocalis actuaría conjuntamente con el cricotiroides para controlar la frecuencia fundamental en el registro de pecho o bajo. La diferencia en el control muscular entre los dos registros provocará una diferencia en el modo de vibración de las cuerdas vocales.

La tensión de las estructuras laringeas varía, asimismo, en función de la activación del organismo, en respuesta a los resultados de las valoraciones realizadas con respecto a sus metas y necesidades. Los estudios con vocalizaciones animales y humanas parecen estar de acuerdo en la existencia de un incremento de la frecuencia fundamental y de la energía, en función de tal activación.

Por último, otra dimensión a lo largo de la cual se dan transiciones continuas está relacionada con la valoración del placer intrínseco del estímulo. En respuesta a estímulos agradables o desagradables se produce una mayor o menor apertura del tracto vocal, cuyos resultados acústicos no pueden ser descritos con precisión en este trabajo.

Estas dimensiones acústicas se perfilan en base a las dimensiones de potencia, actividad y evaluación que han sido propuestas por diversos teóricos de la emoción (Harlow y Stagner, 1933; Nowlis y Nowlis, 1956; Plutchick, 1962; Scholsberg, 1954).

A la vista de los resultados recogidos anteriormente, creemos conveniente plantear una serie de cuestiones que orienten teórica y metodológicamente los trabajos a desarrollar en esta área.

En primer lugar consideramos que una gran parte de los estudios revisados carecen de una base teórica sólida que los sustente, por lo que las investigaciones realizadas poseen, a nuestro juicio, un carácter excesivamente inferencialista. La razón estriba en que, a diferencia de lo que ocurre con otras conductas expresivas, las estructuras y procesos fisiológicos que determinan los parámetros acústicos de las vocalizaciones son complejos y aún no pueden ser descritos con precisión; asimismo, el análisis de la conducta vocal no verbal debe de ser contemplado desde diversas disciplinas, lo que dificulta la elaboración de conceptos y métodos de estudio; sin olvidar la dificultad que supone la delimitación de los componentes verbales y no verbales en dicha conducta. Sería, por tanto, de gran interés para el desarrollo de esta área, dirigir parte de nuestro esfuerzo a la consecución de una buena base teórica que

permita explicar adecuadamente nuestros resultados.

Parece necesario, igualmente, perfeccionar los sistemas de evaluación y medida; en este sentido nuestro equipo junto con otros grupos europeos está desarrollando un instrumento de medida de la conducta vocal, a fin de que las mediciones a realizar en futuros trabajos posean una mayor uniformidad. Dicho instrumento se desarrolla como una parte del proyecto europeo VOX. Consideramos, asimismo relevante una aproximación al estudio de la conducta vocal desde una perspectiva multivariada, en línea con el trabajo iniciado por Muñoz (1986), ya que, según los resultados obtenidos en el mismo, la transmisión de la información sobre el estado afectivo del hablante parece depender de la actuación conjunta de determinadas claves no verbales.

Por último, destacar la importancia que, para la generalización de los resultados, tendría la realización de investigación en situaciones naturales, salvando las dificultades técnicas y metodológicas que ello conlleva.

Creemos, pues, que la toma en consideración de tales cuestiones nos permitirá avanzar adecuadamente en el estudio de la expresión y reconocimientos de las emociones a través de la conducta vocal.

Referencias

- Abercrombie, D. (1968): Paralanguage, *British Journal of Disorders of Communication*, 3, 55-59.
- Albas, D. C.; McCluskey, K. W., y Albas, C. A. (1976): Perception of the emotional content of speech: A comparison of two Canadian groups, *Journal of Cross-cultural Psychology*, 7, 481-490.
- Apple, W., y Hecht, K. (1982): Speaking emotionally: The relation between verbal and vocal communication of affect, *Journal of Personality and Social Psychology*, 42 (5), 864-875.
- Bortz, J. (1966): Physikalisch-akustische korrelate der vocalen kommunikation, *Arbeiten aus dem Psychologischen Institut der Universitate Hamburg*, 9.
- Collier, G. (1985): *Emotional Expression*, London, Lawrence Erlbaum Publishers.
- Constanzo, F. S., Markel, N. N., y Constanzo, P. R. (1969): Voice quality profile and perceived emotion, *Journal of Counseling Psychology*, 16, 267-270.
- Davitz, J. R. (1964): *The Communication of Emotional Meaning*, New York, McGraw-Hill.
- Davitz, J. R., y Davitz, L. J. (1959): The communication of feelings by content-free speech, *Journal of Communication*, 9, 6-13.
- Del Pozo, F.; Martinez, R.; Jimenez, A.; Zaplana, J., y Delgado, J. M. (1981): Objetivaci3n no verbal del lenguaje hablado, *Simposium Nacional de Ingenieria Biomedica*, 1.
- Dimitrovsky, L. (1964): The ability to identify the emotional meaning of vocal expressions at successive levels. En J. R. Davitz (ed.): *The Communication of Emotional Meaning* (pp. 69-86), New York, McGraw-Hill.
- Duncan, S. D., y Rosenthal, R. (1968): Vocal emphasis on experimenters' instruction reading as unintended determinant of subjects' responses, *Language and Speech*, 2, 20-26.
- Dusenbury, D., y Knower, F. H. (1938): Experimental studies of the symbolism of action and voice (III): A study of the specificity of meaning in facial expression, *Quarterly Journal of Speech*, 24, 424-435.
- Eldred, S. H., y Price, D. B. (1958): A linguistic evaluation of feeling states in psychotherapy, *Psychiatry*, 21, 115-121.
- 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.
- Fairbanks, G., y Pronovost, W. (1939): An experimental study of the pitch characteristics of the voice during the expression of the emotion, *Speech Monographs*, 6, 87-104.
- Feldstein, S.; Jaffe, J., y Cassotta, L. (1968): The discrimination and identification of simulated affective vocal behavior, *Research Bulletin*, 3, 1-28.
- F3nagy, I., y Magdics, K. (1963): Emotional patterns in intonation and music, *Phonetica*, 16, 293-326.
- Fry, D. B. (1976): Voice and speech. En R. Hinchcliffe y D. Harrison (eds.): *Scientific Foundations of Otolaryngology*, London, Heineman Medical Books.
- Green, R., y Cliff, N. (1975): Multidimensional comparisons of structures of vocally and facially expressed emotion, *Perception and Psychophysics*, 17, 429-438.
- Harlow, H. F., y Stagner, R. (1933): Psychology of feelings and emotions, II. Theory of emotions, *Psychological Review*, p 191.
- Huttar, G. L. (1968): Relations between prosodic variables and emotions in normal American English utterances, *Journal of Speech and Hearing Research*, 11, 481-487.
- Jimenez, A. (1987): *Marcadores emocionales en la conducta vocal*, Madrid, Universidad Aut3noma de Madrid.
- Kaiser, L. (1962): Communication of affects by single-vowels, *Synthese*, 14, 300-319.
- Knower, F. H. (1941): Analysis of some experimental variations of simulated vocal expressions of the emotions. *Journal of Social Psychology*, 14, 369-372.
- Kramer, E. (1964): Elimination of verbal cues in judgments of emotion from voice, *Journal of Abnormal and Social Psychology*, 64, 390-396.
- Krauss, R.; Apple, W.; Morency, N.; Wenzel, C., y Winton, W. (1981): Verbal, vocal and visible factors in judgments of another's affect, *Journal of Personality and Social Psychology*, 40, 312-320.
- Levin, H., y Lord, W. (1975): Speech pitch frequency as an emotional state indicator, *IEEE Transactions and Systems, Man and Cybernetics Smc-5*, 2, 259-272.
- Lieberman, P., y Michaels, S. B. (1962): Some aspects of fundamental frequency and envelope amplitude as related to the emotional content of speech, *Journal of the Acoustical Society of America*, 34, 922-927.
- Lindsley, D. B. (1951): Emotion. En S. S. Stevens (ed.): *Handbook of Experimental Psychology* (pp. 473-516), New York, Wiley.
- Mallo, M. J. (1986): *Conducta vocal y reconocimiento de emociones*, Memoria de Licenciatura, Universidad Aut3noma de Madrid, Madrid.
- Markel, N. N. (1965): The reliability of coding paralanguage: Pitch, loudness, and tempo, *Journal of Verbal Learning and Verbal Behavior*, 4, 306-308.
- Markel, N. N.; Bein, M. F., y Phillis, J. A. (1973): The relationship between words and tone of voice, *Language and Speech*, 16, 15-21.
- McCluskey, K. W.; Niemi, R. R., y Albas, D. C. (1978): Vocal communication of emotional meaning among normal and disturbed children, *Journal of Special Education*, 12, 443-449.

- Mehrabian, A. (1972): *Nonverbal Communication*, Chicago, Aldin.
- Muñoz, C. (1986): *Reconocimiento de patrones en la expresión emocional a través de la conducta vocal*, Memoria de Licenciatura, Universidad Autónoma de Madrid, Madrid.
- Nowlis, V., y Nowlis, H. H. (1956): The description and analysis of moods, *Annals of the New York Academy of Science*, 65, 345-355.
- Osgood, Ch.; Suci, G. J., y Tannenbaum, P. H. (1976): *La medida del significado*, Madrid, Gredos.
- Pfaff, P. L. (1954): An experimental study of the communication of feeling without contextual material, *Speech Monographs*, 21, 155-156.
- Plutchick, R. (1962): *The emotions: Facts, Theories and a New Model*, Nueva York, Random House.
- Pollack, I.; Rubenstein, H., y Horowitz, A. (1960): Communication of verbal modes of expression, *Language and Speech*, 3, 121-130.
- Ross, M.; Duffy, R. J., y Cooker, H. S. (1973): Contribution of the lower audible frequencies to the recognition of emotions, *American Annals of the Deaf*, 118, 37-42.
- Scherer, K. R. (1971): Randomized splicing: A note on a simple technique for masking speech content, *Journal of Research in Personality*, 5, 155-159.
- Scherer, K. R. (1974): Acoustic concomitants of emotional dimensions judging affect from synthesized tone sequences. En S. Weitz (ed.): *Nonverbal Communication*, Oxford, Oxford University Press.
- Scherer, K. R. (1978): Personality inference from voice quality: The loud voice of extraversion, *European Journal of Social Psychology*, 8, 467-487.
- Scherer, K. R. (1981): Speech and emotional states. En J. K. Darby (ed.): *Speech Evaluation in Psychiatry* (pp. 189-220), New York, Grune and Stratton.
- Scherer, K. R. (1984): On the nature and function of emotion: A component process approach. En K. R. Scherer y P. Ekman (eds.): *Approaches to Emotion*, Hillsdale, N. J., Erlbaum.
- Scherer, K. R., y Giles, H. (eds.) (1979): *Social Markers in Speech*, Cambridge, Cambridge University Press.
- Scherer, K. R., Koivumaki, J., y Rosenthal, R. (1972): Minimal cues in the vocal communication of affect, *Journal of Psycholinguistic Research*, 1, 269-285.
- Scherer, K. R.; Ladd, D. R., y Silverman, K. E. (1984): Vocal cues to speaker affect: Testing two models, *The Journal of the Acoustical Society of America*, 76, 1346-1356.
- Scherer, K. R.; London, H., y Wolf, J. (1973): The voice of confidence: Paralinguistic cues and audience evaluation, *Journal of Research in Personality*, 7, 31-44.
- Scherer, K. R., y Oshinsky, J. S. (1977): Cue utilization in emotion attribution from auditory stimuli, *Motivation and Emotion*, 1, 331-346.
- Scholsberg, H. (1954): Three dimensions of emotion, *Psychological Review*, 61, 81-88.
- Sedláček, P. V., y Sychra, A. (1963): Die Melodie als Faktor des emotionellen Ausdrucks, *Folia Phoniatica*, 15, 89-98.
- Simonov, P. V.; Frolov, M. V., y Taubkin, V. L. (1975): Use of the invariant method of speech analysis to discern the emotional state of announcers, *Aviation, Space and Environmental Medicine*, 46, 1014-1016.
- Skinner, E. R. (1935): A calibrated recording and analysis of the pitch, force and quality of vocal tones expressing happiness and sadness and a determination of the pitch and force of the subjective concepts of ordinary, soft and loud tones, *Speech Monographs*, 2, 81-137.
- Soskin, W., y Kauffman, P. (1961): Judgment of emotion in word free voice samples, *Journal of Communication*, 11, 73-80.
- Van Bezoyen, R. (1984): *Characteristics and Recognizability of Vocal Expressions of Emotion*, Dordrecht-Holland, Foris Publications.
- Williams, C. E., y Stevens, K. N. (1972): Emotion and speech: Some acoustical correlates, *The Journal of the Acoustical Society of America*, 52, 1238-1250.
- Williams, C. E., y Stevens, K. N. (1981): Vocal correlates of emotional states. En J. K. Darby (ed.): *Speech Evaluation in Psychiatry* (pp. 221-240), New York, Grune and Stratton, 221-240.
- Zuberbier, E. (1957): Zur Schreib- und Sprechmotorik der Depression, *Zeitschrift für Psychotherapie und Medizinische Psychologie*, 7, 239-249.
- Zuckerman, M.; Hall, J. A.; De Frank, R. S., y Rosenthal, R. (1976): Encoding and decoding of spontaneous and posed facial expressions, *Journal of Personality and Social Psychology*, 34, 966-977.