

# El reconocimiento de emociones a través de la voz\*

M.<sup>a</sup> José Mallo Carrera

Alfonso Jiménez Fernández\*\*

El presente artículo constituye un intento de evaluar la exactitud con que los sujetos pueden reconocer cuatro estados emocionales básicos: alegría, tristeza, miedo y cólera, a partir de la información vocal no verbal del habla. Para ello los sujetos fueron expuestos a una misma frase de contenido neutro repetida por un actor profesional simulando los cuatro estados emocionales básicos y un estado neutro. Se pretende también evaluar la contribución relativa de un parámetro acústico concreto el perfil de tono en el reconocimiento emocional a través de la presentación frases neutras cuyos valores en el perfil de tono fueron artificialmente aproximados a los de las frases juzgadas como emocionales mediante la utilización de técnicas de procesamiento y síntesis digital de voz (modelo LPC). Los resultados indican que los rasgos vocales no verbales del habla transmiten la suficiente información como para permitir altos porcentajes de acierto en la identificación de las cuatro categorías emocionales manejadas. Asimismo, los sujetos fueron capaces de reconocer las emociones de tristeza y cólera a partir únicamente de sus perfiles de tono característicos.

La investigación sobre la expresión emocional ha estado prioritariamente centrada en el canal facial, y aunque ya en 1872 el propio Darwin señaló la importancia de los aspectos no verbales del habla como medio de expresión, su estudio sistemático sólo ha tenido lugar en la última década. Las dificultades de tipo técnico que conlleva la evaluación de la conducta vocal y la confusión respecto a su estatus como conducta no verbal son dos de las razones que han motivado este desfase con respecto a la expresión facial (Scherer, 1982).

Los estudios sobre expresión vocal de la emoción pueden encuadrarse en dos grandes grupos: los estudios de «expresión» y los de «reconocimiento». Los primeros, también llamados de «externalización» (Brown y Bradshaw, 1985) tienen como objetivo básico determinar cómo un estado emocional se expresa o exterioriza a través de la voz del sujeto. La evaluación de ésta puede realizarse a varios niveles, entre ellos, el análisis acústico y los juicios perceptivos de los parámetros acústicos que han sido los más utilizados. Los trabajos de Fairbanks (Fairbanks y Pronovost, 1939; Fairbanks, 1940; Fairbanks y Hoaglin, 1941), E. R. Skinner (1935), Williams y Stevens (1969, 1972), Levin y Lord (1975), y Scherer y Oshinsky (1977), han constituido algunos de los intentos más relevantes dentro de esta línea.

Los estudios de reconocimiento tratan de conocer en qué medida el receptor es capaz de identificar, a través de los aspectos no verbales de la voz

\* Esta investigación ha sido en parte financiada con cargo al proyecto N.º 2638/83.C2 de la Comisión Asesora para la Investigación Científica y Técnica (CAYCIT).

\*\* Departamento de Psicología Básica, Social y Metodología (Universidad Autónoma de Madrid).

del emisor, el estado emocional expresado por éste, de ahí que se denominen también estudios de «exactitud» (Rosenthal, 1982; Brown y Bradshaw, 1985). Uno de los principales problemas con que se ha enfrentado la investigación sobre reconocimiento ha sido la separación del canal verbal del vocal, de modo que el contenido lingüístico del mensaje no condicione los juicios sobre el estado afectivo del codificador. Aunque han sido varios los métodos utilizados para controlar la información verbal: filtrado de frecuencias (Soskin y Kaufmann, 1964; Burns y Beier, 1973; Fechner, 1978; McCluskey, Albas y Niemi, 1978; Scherer, Koivumaki y Rosenthal, 1972; Scherer, Ladd y Silverman, 1984), «random splicing» (Blanck, Rosenthal, Snodgrass, DePaulo y Zuckermann, 1981; Scherer, *et al.*, 1972; Scherer *et al.*, 1984) la opción más aceptada ha consistido en mantener constante una frase con contenido neutro para todas las emociones (Fairbanks y Pronovost, 1939; Zuckermann, Lipets, Koivumaki, 1975; Fenster, Blake y Goldstein, 1977; Williams y Stevens, 1981; Van Bezooijen, 1984).

Dentro de esta segunda línea de investigación se encuadran los trabajos que vamos a presentar a continuación. El primero de ellos constituye un estudio clásico de reconocimiento en tanto que pretende evaluar la exactitud con que cuatro estados emocionales básicos (alegría, tristeza, miedo y cólera) pueden ser reconocidos por los sujetos, cuando éstos únicamente disponen de la información vocal no verbal. Junto a este objetivo prioritario, estábamos interesados en otras cuestiones como el diferente grado de reconocimiento entre las categorías emocionales y el tipo de confusiones que se producen entre ellas, cuestiones que, por desgracia, no todos los estudios de reconocimiento han abordado.

Casi todos los trabajos informan sobre el hecho de que unas emociones son mejor reconocidas que otras; sin embargo, el problema reside en que los resultados son muy diversos en cuanto a qué emociones en concreto obtienen los mayores porcentajes de elecciones correctas. La hipótesis más barajada para explicar las diferencias en los grados de reconocimiento de las distintas emociones postula que las emociones negativas son mejor identificadas que las positivas. Este punto de vista es apoyado por los estudios evolutivos que han investigado esta cuestión (Dimitrosky, 1964; Fenster *et al.*, 1967; Mac Cluskey, Niemi y Albas, 1978). En concreto, Fenster plantea que el mejor reconocimiento de las emociones negativas puede explicarse en términos de su mayor «valor de supervivencia», en el sentido de que un error en la decodificación de expresiones que reflejan emociones negativas puede tener repercusiones más graves que los errores en la decodificación de emociones positivas.

En los estudios con adultos la cuestión no está tan clara. Scherer (1981) considera, tras su revisión de estudios de reconocimiento, que la hipótesis puede mantenerse y señala que cólera suele ser la emoción mejor reconocida, seguida de tristeza, indiferencia y alegría. Sin embargo, existen resultados contradictorios como los de Brighetti, Ladavas, y Ricci Bitti (1980) que refieren que con excepción de tristeza, las emociones positivas (alegría y sorpresa) fueron mejor reconocidas que las negativas (miedo, cólera, asco y desprecio).

El análisis de las confusiones constituye otro de los principales temas de interés en los estudios de reconocimiento, postulándose dos dimensiones básicas capaces de explicar los errores que cometen los sujetos: la dimensión evaluativa y la de activación. Las conclusiones de los estudios de reconocimiento revisados (Fairbanks y Pronovost, 1939; Davitz y Davitz, 1959; Kramer, 1964; Feldstein, Jaffe y Cassotta, 1968), incluyendo los evolutivos (Dimitrosky, 1964; Van Bezooijen, 1984) y los transculturales (Albas, McCluskey y Albas, 1976; Van Bezooijen, 1984) dan un mayor peso a la dimensión ac-

tivación en la categorización de los estímulos emocionales, proponiendo dicha dimensión como la responsable de las confusiones entre los distintos estados afectivos.

Si esta hipótesis es cierta, cabría esperar, en el presente experimento, que las emociones activas (alegría y cólera) fueran confundidas entre sí y que lo mismo ocurriera entre las emociones pasivas (tristeza y miedo). Por el contrario, las confusiones entre cólera y tristeza hablarían a favor de la proximidad en el significado evaluativo como causa de los errores. No obstante, debe considerarse que en nuestro caso resulta difícil probar el papel de la dimensión evaluación en la discriminación entre estados afectivos, en la medida en que sólo se cuenta con una emoción positiva frente a tres negativas.

En el segundo experimento se pretende estudiar la contribución relativa de un parámetro acústico concreto, el perfil de tono, al reconocimiento de emociones. Esta perspectiva resulta mucho menos clásica ya que han sido muy pocos los trabajos que han tratado de relacionar los resultados de los análisis acústicos de las expresiones emocionales con la decodificación de éstas por parte de los oyentes. Lo que se pretende, dentro de esta línea, es descubrir cuáles son las señales objetivas de la conducta vocal del emisor que conducen a la atribución de un determinado estado emocional por parte del decodificador. La mayor parte de los trabajos con este objetivo han utilizado una metodología correlacional que mediante técnicas como la regresión múltiple permiten conocer qué parámetros acústicos explican la mayor parte de la varianza de los juicios (Scherer y Oshnisky, 1977; Mallard y Daniloff, 1979). La aplicación de la metodología experimental exige la manipulación controlada de la conducta vocal (Vi) y la observación del efecto de dicha manipulación sobre los juicios de los sujetos (Vd). Lógicamente, ello exige contar con dispositivos instrumentales y analíticos altamente sofisticados y ésta, concretamente, ha sido la razón por la cual la investigación en este campo es apenas incipiente.

En un trabajo anterior de uno de los autores (Jiménez Fernández, 1985), la manipulación de la conducta vocal fue posible gracias al modelo de predicción lineal (LPC) de análisis/síntesis que permite modificar uno o varios parámetros asignándoles aquellos valores que se deseen y sintetizar de nuevo las muestras de habla modificadas. De esta forma, es posible examinar los efectos que tiene sobre el reconocimiento la manipulación de cualquier parámetro aislado. Los parámetros que pueden ser objeto de modificación son innumerables, por ello la elección de la variable a manipular debe estar guiada por algún criterio que lo señale como relevante. En este sentido se eligió la frecuencia fundamental, dado que por su generación glótica parece estar especialmente afectada por el «arousal» emocional (Scherer, 1979), asimismo son varios los trabajos anteriores que indican su importancia para discriminar entre estados emocionales (Williams y Stevens, 1972; Levin y Lord, 1975; Hauser, 1976; Streeter, Krauss, Apple y MacDonald, 1978). Por otra parte los propios datos descriptivos aportados por el análisis acústico de las expresiones sin modificar, realizado en el citado estudio (Jiménez Fernández, 1985), mostraron que los parámetros de tono discriminaban significativamente entre las emociones consideradas.

Por ello, las frases originalmente neutras fueron aproximadas en los parámetros de frecuencia fundamental a los valores medios que caracterizaron en el estudio analítico a las frases reconocidas como emocionales: alegría, tristeza, miedo y cólera. La hipótesis barajada consistió en considerar que si la frecuencia fundamental es el parámetro más relevante en la identificación de la expresión emocional, las frases neutras modificadas serían correctamente reconocidas por los sujetos como representantes de los distintos estados emo-

cionales. Sin embargo, los sujetos que previamente habían identificado de modo correcto las expresiones sin modificar no fueron capaces de reconocer las frases modificadas por encima del grado esperado por azar, hecho que no podía ser atribuido a la calidad de audición de las frases modificadas que fue calificada de «buena», sobre una escala de cuatro puntos (0 a 3) correspondientes a las etiquetas: «muy mala», «mala», «buena» y «muy buena».

Antes de cuestionar el papel de la Fo en el reconocimiento de emociones, nos planteamos replicar el trabajo introduciendo algunas modificaciones. En el estudio que va a replicarse las respuestas de los jueces ante las frases modificadas fueron recogidas utilizando un formato con seis alternativas, mientras que las categorías estimulares presentadas eran únicamente cuatro: las correspondientes a los estados de alegría, tristeza, miedo y cólera. Además de ellas, el sujeto tenía la posibilidad de elegir la categoría «No reconoce» y «Neutra», a pesar de que esta última no fue presentada como estímulo. Por ello, y a la vista de los resultados donde el mayor porcentaje de respuestas se acumulaba en la categoría neutra, nos planteamos la posibilidad de replicar el trabajo modificando el sistema de elección. Se pensó, por tanto, en eliminar las alternativas de respuesta «Neutra» y «No reconoce», forzando, así, a los sujetos a elegir entre las categorías que representan opciones de respuesta más conservadoras. Por otra parte, algunos autores (Van Bezooijen, 1984) han señalado que es posible que el estado neutro merezca una consideración independiente del resto de los estados afectivos en la medida en que la distinción entre estados emocionales y no emocionales podría ser de distinta naturaleza que la discriminación entre las diferentes emociones.

Nuestro objetivo es, por tanto, determinar si con este procedimiento el porcentaje de elecciones correctas se sitúa por encima del esperado por azar, comprobando con ello el papel de la frecuencia fundamental en el reconocimiento de emociones a través de la conducta vocal.

## EXPERIMENTO I

### Método

#### *Sujetos*

En el experimento participaron 90 sujetos, elegidos aleatoriamente entre los alumnos de las facultades de Psicología y Económicas de la Universidad Autónoma de Madrid. La muestra estaba compuesta por 25 varones y 65 mujeres, su rango de edades oscilaba entre dieciocho y treinta y cinco años.

#### *Diseño*

Se utilizó un diseño unifactorial y univariado de medidas repetidas siendo la variable independiente la emoción presentada como estímulo, con cuatro niveles o tratamientos correspondientes a cada uno de los estados emocionales considerados: alegría, tristeza, miedo y cólera. La variable dependiente fue la emoción elegida por el sujeto ante cada uno de los estímulos.

El material estimular empleado fue el mismo utilizado en los experimentos a los que se pretende replicar (Jiménez Fernández, 1985). Las expresiones originales se obtuvieron con la colaboración de un actor profesional, don José María Rodero, que simuló a través de la frase «El papel tiene un peso neto» los distintos estados emocionales. El valor neutro de la frase se probó

previamente mediante la técnica del diferencial semántico. La frase fue repetida seis veces para cada uno de los tratamientos; dado que en el presente estudio no se presentó la categoría neutra, los sujetos fueron expuestos a un total de 24 estímulos (4 emociones × 6 frases por emoción).

### *Instrumentación*

Para la presentación de los estímulos se utilizó un equipo de reproducción acústica de alta fidelidad, compuesto por un magnetófono TANDBERG TD20A, un micrófono marca AKAI y un altavoz.

Los juicios de los sujetos fueron recogidos en una hoja de respuestas, donde para cada estímulo debía señalarse una de las cuatro emociones ofrecidas. El análisis de los datos se realizó con ayuda del ordenador VAX11/780 del Centro de Cálculo de la U. A. M.

### *Procedimiento*

El experimento tuvo lugar en el laboratorio de conducta vocal del área de Psicología Social que dispone de una sala insonorizada donde fueron acomodados los sujetos y de una cabina adyacente de control en la que se situó el experimentador. Desde la cabina es posible el acceso visual a la sala a través de un espejo unidireccional, al mismo tiempo el experimentador podía comunicarse desde la cabina con los sujetos mediante un micrófono conectado con la sala.

A continuación adjuntamos las instrucciones que recibieron los sujetos; éstas figuraban en la hoja de respuestas, pero fueron, además, leídas por el experimentador antes de dar comienzo a la presentación de los estímulos:

Van a escuchar una frase repetida 24 veces por el mismo sujeto, cada vez reflejará una emoción (alegría, tristeza, miedo y cólera). Tienen que marcar con una cruz la casilla correspondiente a la emoción con que, a su juicio, se emitió cada frase. En caso de duda, trate de responder eligiendo aquella emoción a la que mejor se ajuste la frase pero, por favor, no conteste al azar.

Escucharán una a una las 24 frases que tendrán siempre el mismo contenido: «El papel tiene un peso neto». Escriban, ahora, su nombre, edad, curso y marquen la casilla correspondiente a su sexo.

Si tienen alguna duda, consúltenla antes de comenzar.

La presentación de los estímulos se realizó de forma colectiva, en grupos de diez sujetos. Para evitar los efectos de la posición de los estímulos dentro de la serie sobre las respuestas, los 24 estímulos se dividieron en tres bloques de ocho, dentro de cada bloque el orden de presentación fue aleatorizado, obteniéndose las siguientes secuencias:

Bloque 1: T6, A5, C1, A1, C6, M6, T3, T1  
Bloque 2: C3, T4, A3, M5, M3, C4, T5, C2  
Bloque 3: M1, C5, M4, T2, A2, M2, A4, A6

El orden de los bloques fue contrabalanceado para los nueve grupos siguiendo un diseño cuadrado latino, de tal forma que cada uno de los bloques es presentado el mismo número de veces en posición inicial, final y media.

(1, 2, 3)    (3, 2, 1)    (2, 3, 1)  
(2, 3, 1)    (1, 2, 3)    (3, 1, 2)  
(3, 1, 2)    (2, 3, 1)    (1, 2, 3)

## Resultados

La cuantificación de las respuestas de los sujetos se realizó siguiendo el modelo de dos pasos, propuesto por Rosenthal (1982), que constituye la forma más habitual de trabajar con los datos, cuando se utilizan formatos categoriales. Dicho modelo permite no sólo obtener el número de elecciones correctas para cada emoción (diagonal de la matriz), sino también la matriz completa de puntuaciones, lo cual posibilita el estudio de la dirección de los errores. Aplicando este modelo se obtuvieron las matrices de datos para los 90 jueces que aparecen representados en la tabla I, donde las puntuaciones son ofrecidas en valores absolutos y en proporciones.

TABLA I

*Matriz completa de datos para todos los jueces empleando frecuencias absolutas (tabla superior) y frecuencias relativas expresadas en proporciones (tabla inferior)*

Categoría elegida	Categoría correcta				Total
	Alegría	Tristeza	Miedo	Cólera	
Alegría	382	8	26	10	426
Tristeza	55	433	77	7	572
Miedo	39	83	414	31	567
Cólera	59	13	19	491	582
No contesta	5	3	4	1	13
Total	540	540	540	540	2.160

Categoría elegida	Categoría correcta				Total
	Alegría	Tristeza	Miedo	Cólera	
Alegría	0.707	0.015	0.048	0.018	0.788
Tristeza	0.102	0.802	0.142	0.013	1.059
Miedo	0.072	0.154	0.767	0.057	1.050
Cólera	0.109	0.024	0.035	0.909	1.077
No contesta	0.009	0.005	0.007	0.002	0.024
Total	1.000	1.000	1.000	1.000	4.000

Antes de proceder a la realización de los análisis inferenciales se calculó la fiabilidad media entre jueces. Dado que en este estudio han sido empleados 90 sujetos, el cálculo de la fiabilidad efectiva (R) en función del coeficiente de correlación medio entre pares de jueces resulta muy costoso (Rosenthal, 1982). Por ello, se utilizó la técnica de Guilford que permite estimar la fiabilidad efectiva a partir del análisis de varianza de dos factores (jueces/juicios). En la tabla II pueden verse las estimaciones de fiabilidad para el conjunto total de estímulos y para cada una de las categorías emocionales.

Examinaremos, a continuación, los resultados relativos a las cuestiones que se plantearon en la introducción.

TABLA II

*Estimación de los índices de fiabilidad (efectivo: R y fiabilidad media: r) entre los 90 jueces en sus valoraciones de reconocimiento de las frases originales (por emociones y para el total del experimento)*

Fuente	SC	GL	MS	r (est.)	R (est.)
Juicios	37.10370	5	7.42074	0.3769	0.9819
Jueces	15.10370	89	0.16970		
Residuo	59.56296	445	0.13385	Alegría	
Juicios	9.72037	5	1.94407	0.1268	0.9289
Jueces	14.63148	89	0.16440		
Residuo	61.44630	445	0.13808	Tristeza	
Juicios	3.28899	5	0.65778	0.0317	0.7470
Jueces	19.26667	89	0.21648		
Residuo	74.04444	445	0.16639	Miedo	
Juicios	1.92037	5	0.38407	0.0439	0.8055
Jueces	9.38704	89	0.10547		
Residuo	33.24630	445	0.07471	Cólera	
Juicios	63.68148	23	2.76876	0.1890	0.9539
Jueces	25.95370	89	0.29161		
Residuo	260.73519	2047	0.12737	Total	experimento

1. *¿Son reconocidas las emociones por encima del grado esperado por azar?*

Para probar la hipótesis del reconocimiento más allá del azar se utilizó la media de frases reconocidas para cada tratamiento y para el conjunto total del experimento. Esta media, que puede oscilar entre 0 y 6, se comparó con la esperada por azar, 1.5 frases acertadas, dado que con cuatro categorías de respuesta la probabilidad de acertar un ítem es de 0.25, mediante la prueba *t* de Student. La comparación resultó significativa (ver tabla III) para todas las emociones: alegría ( $t(89)=25.79$ ,  $p<.01$ ), tristeza ( $t(89)=31.65$ ,  $p<.01$ ), miedo ( $t(89)=25.80$ ,  $p<.01$ ), cólera ( $t(89)=47.15$ ,  $p<.01$ ) y, lógicamente, para la media total del experimento ( $t(89)=28.74$ ,  $p<.01$ ).

TABLA III

*Media de frases reconocidas y desviaciones típicas. El asterisco indica qué valores resultaron estadísticamente significativos ( $p < .01$ )*

Emoción	Media	Desviación típica
Alegría	4.244*	1.009
Tristeza	4.811*	0.993
Miedo	4.600*	1.140
Cólera	5.456*	0.796
Total	4.778*	1.081

Evidentemente, en lugar de la media podríamos haber utilizado los porcentajes de aciertos comparándolos con el porcentaje esperado por azar (25 por 100) mediante un contraste sobre la proporción. Dichos porcentajes pueden ser deducidos a partir de la diagonal de la matriz completa de datos y aparecen redondeados en la primera columna de la tabla IV.

TABLA IV

*Porcentaje de elecciones correctas para cada categoría emocional, incluyendo los resultados tras la corrección por azar, la corrección por tendencia de respuesta y el porcentaje total de elecciones. El número entre paréntesis indica el rango, desde 1 (emoción mejor reconocida) hasta 4 (emoción peor reconocida)*

Emoción	Elecciones correctas (%)	Corrección por azar (%)	Total de respuestas (%)	Corrección tendencia de respuesta (%)
Alegría	70 (4)	61	79	80 (2)
Tristeza	80 (2)	73	106	78 (3)
Miedo	77 (3)	69	105	75 (4)
Cólera	91 (1)	88	108	87 (1)

Utilizamos estos porcentajes para comparar los resultados obtenidos en este experimento con los del trabajo anterior (ver tabla VIII). Dado que el número de alternativas de respuesta no coincide con ambos estudios, la comparación de los resultados sólo resulta posible si los porcentajes son previamente corregidos (Rosenthal, 1982; Van Bezooijen, 1984). Los resultados de la corrección por azar desaparecen en la columna 2 de las tablas IV y V.

TABLA V

*Porcentaje de elecciones correctas obtenido por Jiménez Fernández (1985) incluyendo los resultados tras la corrección por azar, la corrección por tendencia de respuesta y el porcentaje total de elecciones. El número entre paréntesis indica el rango, desde 1 (emoción mejor reconocida) hasta 4 (emoción peor reconocida)*

Emoción	Elecciones correctas (%)	Corrección por azar (%)	Total de respuestas (%)	Corrección tendencia de respuesta (%)
Alegría	65 (5)	58	80	73 (2)
Tristeza	72 (3)	67	111	69 (3)
Miedo	75 (1)	71	122	68 (4)
Cólera	74 (2)	69	89	78 (1)
Neutra	68 (4)	61	98	68 (4)

Como puede observarse (ver figura 1), la eliminación de la categoría neutra produjo un incremento de los porcentajes de elecciones correctas para las emociones de alegría, tristeza y cólera; sin embargo, sólo en el caso de cólera el incremento alcanzó la magnitud suficiente para ser considerado estadísticamente significativo: alegría ( $Z=0.43$ ,  $p>.05$ ), tristeza ( $Z=0.9$ ,  $p>.05$ ) y cólera ( $Z=4.9$ ,  $p<.01$ ). En miedo se produjo una ligera disminución de la puntuación de reconocimiento que tampoco resultó significativa ( $Z=0.70$ ,  $p>.05$ ).



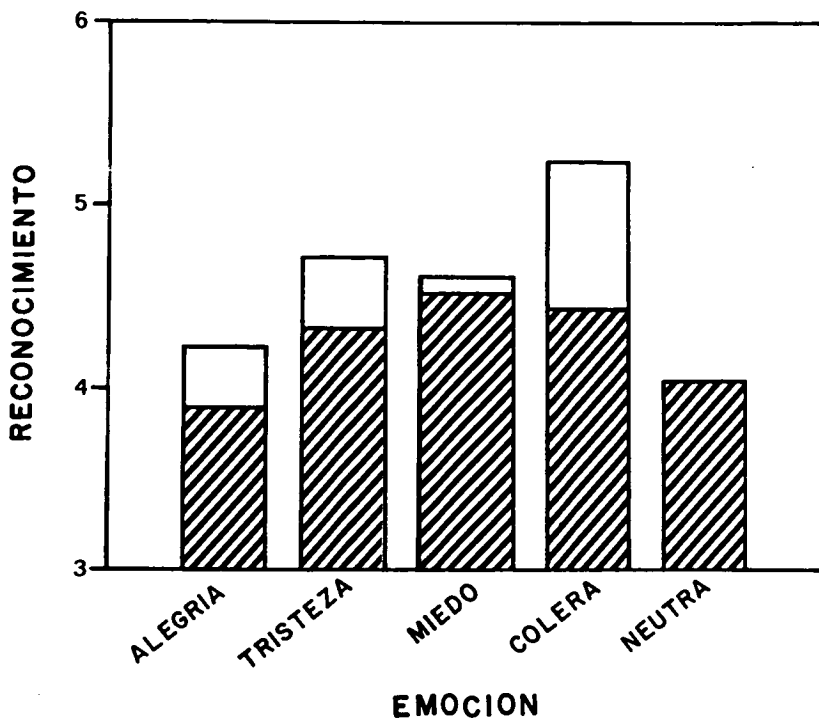


FIGURA 1

Comparación de las puntuaciones medias de reconocimiento obtenidas en este trabajo con las obtenidas por Jiménez Fernández, 1985 (área rayada).

## 2. ¿Son unas emociones mejor reconocidas que otras?

Aunque todas las frases originales fueron reconocidas por encima del azar, la simple inspección de los porcentajes de elecciones correctas (cólera: 91 por 100, tristeza: 80 por 100; miedo: 77 por 100 y alegría: 70 por 100) o la media de frases reconocidas (cólera: 5,45; tristeza: 4,81; miedo: 4,60 y alegría: 4,24) nos pone sobre la pista de la existencia de diferencias entre las emociones respecto a la facilidad con que son reconocidas por los sujetos.

De hecho, los resultados del ANOVA de un factor efectos fijos y medidas repetidas llevado a cabo mediante un programa BMDP2V, revelan la existencia de un efecto significativo del factor emoción ( $F(3) = 32.96, p < .001$ ).

Para determinar qué niveles del factor difieren significativamente de los demás, se aplicó la prueba de comparaciones múltiples de Tukey, adoptando el nivel de significación del 5 por 100. Tras este análisis todas las comparaciones resultaron significativas excepto la comparación tristeza-miedo. De modo que es posible afirmar que cólera fue la emoción mejor reconocida, seguida de tristeza y miedo sin diferencias significativas entre ellas y finalmente alegría, que resultó ser la emoción peor identificada.

Como puede observarse en la columna tercera de la tabla IV, no todas las emociones han sido elegidas con la misma frecuencia: tristeza, miedo y cólera recibieron un porcentaje de respuestas considerablemente mayor que alegría. Para controlar la posibilidad de que el grado de reconocimiento para una categoría aparezca artificialmente «inflado» por la elección aleatoria y sistemática de dicha categoría fue aplicada la corrección por tendencia de respuesta (Rosenthal, 1982; Van Bezooijen, 1984), mostrándose los datos corregidos en la cuarta columna de la tabla IV.

Aun después de la corrección, cólera continúa siendo la emoción mejor reconocida; sin embargo, alegría pasa del último lugar al segundo, seguida de tristeza y miedo con porcentajes de reconocimiento ligeramente inferiores. Los datos obtenidos por Jiménez Fernández (1985) fueron también sometidos a la corrección por tendencia de respuesta. Curiosamente, los rangos de los porcentajes de reconocimiento coinciden en ambos experimentos, una vez que han sido corregidos, pese a que eran algo discrepantes antes de la corrección.

### 3. *Análisis de confusiones*

A este nivel de análisis nos planteamos no sólo explorar la dirección de los errores, sino comprobar qué confusiones se producían con la suficiente frecuencia para alcanzar el nivel de significación estadística. Por ello, sometimos los datos obtenidos a un análisis inferencial.

Si los sujetos al hacer una elección errónea no tienen una preferencia particular por alguna de las categorías restantes, cabría esperar una distribución aleatoria del número de errores entre las alternativas no correctas. Si, por el contrario, existen alguna o varias dimensiones que expliquen las confusiones, se elegirán preferentemente determinadas alternativas incorrectas en detrimento de otras.

A nivel operacional, probar esta hipótesis implica comparar el porcentaje de errores esperado por azar en cada categoría, con el porcentaje empírico obtenido y determinar si este último supera al primero de modo significativo. El porcentaje esperado por azar se obtiene dividiendo el número total de elecciones incorrectas que se han producido en cada emoción por el número de alternativas incorrectas posibles y, lógicamente, transformando el resultado en un tanto por ciento.

En la figura 2 se representan, mediante un diagrama de flechas, aquellas confusiones entre emociones cuya frecuencia resultó significativa. Las flechas dobles indican confusiones simétricas, lo que significa que las emociones que comparten una flecha de este tipo han sido mutuamente confundidas entre sí. Las confusiones asimétricas se representan con una flecha simple, situándose la base de ésta en la categoría que hubiera sido correcta, mientras que la punta de flecha se dirige a la categoría erróneamente elegida.

Como puede observarse alegría ha sido excluida de la gráfica, lo que indicaría que los estímulos de esta emoción no han sido erróneamente juzgados como representantes de ninguna otra categoría emocional en concreto y que los estímulos de otras categorías tampoco han sido considerados como pertenecientes a alegría. Por otra parte, miedo y tristeza aparecen ligados por una relación de confusión simétrica y, finalmente, las frases de cólera fueron significativamente juzgadas como frases de miedo.

## Discusión

Los resultados obtenidos muestran que todas las emociones estudiadas pueden ser reconocidas con un grado de exactitud que supera al esperado por mero azar. Ello indica que la información transportada a través del canal vocal contiene las suficientes claves como para permitir a los sujetos identificar los distintos estados emocionales.

Este hecho ha sido ya señalado por toda la literatura sobre reconocimiento emocional a través de la voz; en todos los estudios revisados el porcentaje

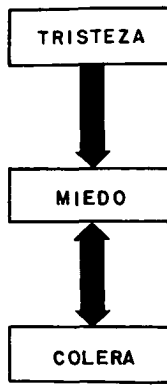


FIGURA 2

*Diagrama de confusiones entre las distintas categorías emocionales.*

de identificaciones correctas es superior a la mera adivinación, incluso en aquellos que refieren los porcentajes de reconocimiento más bajos, como el de Fenster, Blake y Goldstein de 1977 (30 por 100) y el Davitz y Davitz (1959) con un 37 por 100.

Scherer (1981), en su revisión de estudios de reconocimiento calculó la exactitud media alcanzada en dichos estudios, excluyendo del cómputo aquellos que habían utilizado filtrado de frecuencias o lengua extranjera como método de control del contenido verbal o poblaciones especiales de decodificadores. El porcentaje medio obtenido es de un 60 por 100, que se convierte en un 56 por 100 tras aplicarle la corrección por azar. Los resultados de nuestro trabajo superan ampliamente dicha media situándose en un 79.5 por 100 antes de la corrección y en un 72,7 por 100 después de considerar el dato en función del número de alternativas.

La comparación directa con los porcentajes de reconocimiento obtenidos en otros estudios resulta siempre problemática, dada la falta de homogeneidad respecto a varias decisiones metodológicas que, muy probablemente, inciden sobre los resultados. Por ello, quizá resulte de más interés examinar los porcentajes ofrecidos por estudios que, al menos, hayan trabajado con categorías emocionales semejantes a las aquí empleadas. El trabajo de Williams y Stevens (1981) es uno de los más parecidos, no sólo respecto a las emociones estudiadas, que son prácticamente las mismas excepto indiferencia, sino también por el resto de las decisiones metodológicas (contenido verbal constante a través de una frase estándar y actores como codificadores). A pesar de ello, los porcentajes de reconocimiento son distintos; estos autores informan de un 45 por 100 como media de elecciones correctas, frente al 70 por 100 obtenido en nuestro estudio. Y lo mismo sucede respecto a los trabajos de Lewitt, 1964 (47 por 100), Zuckerman et al., 1975 (47 por 100), Fenster et al., 1977 (30 por 100). Resultados más similares a los nuestros son, sin embargo, los ofrecidos por Fairbanks y Pronovost (1939), cuyo trabajo incluye las emociones de tristeza, cólera, miedo, desprecio e indiferencia, con un 79 por 100 de aciertos y el de Kramer (1964) con un 70 por 100.

Respecto a la segunda de las cuestiones planteadas, es claro el hecho de que no todas las emociones son juzgadas con el mismo grado de exactitud. Como puede constatarse en la revisión de la literatura sobre reconocimiento, éste ha sido uno de los temas que más ha preocupado a la investigación en el área. Y, sin embargo, ni los resultados obtenidos son siempre concordantes, ni existen bases teóricas capaces de dar razón de los hallazgos empíricos.

Examinaremos en primer lugar qué relación guardan nuestros datos con los referidos por los trabajos anteriores. Cólera fue la emoción mejor reconocida por los sujetos, resultado bastante acorde con los obtenidos por una gran parte de la investigación: en los estudios de Williams y Stevens (1981), Lewitt (1964), Fenster *et al* (1977), Heike (1969) y Van Bezooijen (1984) ocupa la segunda posición y resultó ser la emoción mejor reconocida en los trabajos de Davitz y Davitz (1959), Bluhme (1971), Burns y Beier (1973) y Ross, Duffy, Cooker y Sargean (1973). De hecho Scherer (1981) indica que cólera es la emoción que obtiene los mayores grados de exactitud en la mayor parte de los estudios revisados por él. Van Bezooijen (1984), sin embargo, considera que cólera es una emoción paradójica en cuanto a las posiciones que ocupa en las clasificaciones de reconocimiento; mientras que en algunos trabajos se sitúa en el primero o segundo lugar, en otros desciende a posiciones medias o medias inferiores.

La emoción de tristeza ocupa la segunda posición, aunque sin diferencias significativas con respecto a miedo. En los estudios revisados, tristeza suele obtener puntuaciones altas de reconocimiento (Williams y Stevens, 1981; Fenster *et al.*, 1977; Zuckermann, 1975; Brighetti *et al.*, 1980), o bien ocupar posiciones intermedias (Dusenbury y Knowler, 1938; Fairbanks y Pronovost, 1939; Kramer, 1964; Van Bezooijen, 1984). Para miedo la situación es más caótica, ocupando en los distintos estudios todo tipo de posiciones.

Respecto a alegría, los datos deben ser interpretados con cautela, ya que aunque los resultados directos la señalan como la emoción peor reconocida, la situación cambia cuando se tiene en cuenta el número total de elecciones recibidas. Tras la corrección por tendencia de respuesta asciende al segundo lugar, situándose por delante de tristeza y miedo. Por lo que respecta a los estudios anteriores esta emoción suele situarse en posiciones medias y medias inferiores, aunque existen también excepciones, como los estudios de Pfaff (1954) y Van Bezooijen (1984) donde aparece entre las emociones mejor reconocidas.

A partir de nuestros datos, resulta bastante complicado determinar si las emociones negativas son mejor reconocidas que las positivas. En primer lugar, la posición de alegría se altera totalmente cuando el porcentaje de reconocimiento se matiza en función del número total de elecciones: alegría resulta ser una emoción poco acertada, pero lo que indica el resultado de la corrección es el hecho de que la mayor parte de las veces que los sujetos han optado por esa categoría en sus respuestas lo han hecho de forma correcta. Por otra parte, en este trabajo, únicamente se utiliza una emoción positiva frente a tres negativas, con lo cual, aunque obviáramos la consideración anterior, resultaría un tanto arriesgado confirmar la hipótesis ante un reparto tan desigual de representantes de ambos tipos de emociones.

El análisis de confusiones fue realizado con el objetivo de determinar qué dimensión de entre las propuestas: evaluación o actividad permite explicar la dirección de los errores. El hecho de que alegría no se confundiera ni fuera confundida con el resto de las emociones, todas ellas negativas, apunta hacia la dimensión evaluativa. Asimismo, la confusión simétrica entre cólera y tristeza y la tendencia a juzgar los estímulos de cólera como miedo, pueden explicarse con cargo a dicha dimensión.

La relación entre tristeza y miedo ya ha sido referida por varios trabajos anteriores: Feldstein *et al* (1968) encontraron que miedo era incorrectamente juzgado como tristeza y Van Bezooijen (1984) informa del mismo error en las decodificaciones de sujetos holandeses y japoneses, mientras que la confusión se producía a la inversa en un grupo de sujetos taiwaneses.

Es cierto, sin embargo, que estas dos emociones, no sólo son similares en

su significado evaluativo, en tanto que se trata de dos estados afectivos negativos, sino que también están próximas respecto a su nivel de actividad. En una investigación paralela (Muñoz, 1986), utilizando estas mismas expresiones, se obtuvieron, a partir de los valores de las frases en los parámetros de  $F_0$  media y dos de los coeficientes de la aproximación parabólica del contorno de tono ( $F_0$ -A1,  $F_0$ -A2), funciones que permitían discriminar las expresiones de tristeza y miedo frente a las de alegría y cólera.

Agrupaciones similares han sido obtenidas utilizando datos perceptivos; Van Bezooijen (1984) realizó un análisis discriminante basándose en los valores que un grupo de jueces atribuyó a diez emociones a lo largo de una serie de dimensiones perceptivas extraídas del protocolo del Laver. En sus resultados tristeza y miedo aparecen de nuevo juntas, formando parte del mismo grupo, llamado «pasivo-negativo», cólera se integra en el «activo-negativo» y, finalmente, alegría está incluida en el grupo «activo-positivo».

Aunque esta confusión simétrica entre tristeza y miedo pueda ser explicada a partir de la dimensión actividad, el resto de los resultados no encajan dentro de esta interpretación. En primer lugar, cabría esperar confusiones entre cólera y alegría que aparecen relacionadas tanto desde un punto de vista acústico como perceptivo. En el trabajo de Van Bezooijen, los jueces atribuyeron a ambas niveles altos y amplios rangos de tono, alto volumen, y ritmo rápido. Los datos referidos por Jiménez Fernández (1985), tras analizar acústicamente las expresiones emocionales que han servido como estímulos en este experimento, indican un incremento de la variabilidad de la  $F_0$ , energía media, variabilidad de ésta y duración, para ambas emociones. Únicamente existen diferencias en el parámetro  $F_0$  media, para el cual cólera no se diferencia de la situación neutra mientras que alegría se sitúa por encima.

Curiosamente, en los estudios de reconocimiento llevados a cabo por Van Bezooijen con el mismo material que analizó en términos perceptivos, aparecen confusiones cólera-alegría; sin embargo, en nuestro estudio, no se presenta dicha confusión. A nivel especulativo podríamos proponer que aunque ambas emociones puedan considerarse activas respecto a la mayoría de los parámetros, su diferente situación en relación a la  $F_0$  media impidió que se produjeran confusiones entre ambas. Recordemos que, sin embargo, en el estudio de Van Bezooijen los jueces atribuyeron a las expresiones de ambas valores altos de nivel de tono, correlato perceptivo de la  $F_0$ .

La tendencia a juzgar erróneamente cólera como miedo tampoco puede explicarse en términos de proximidad en el nivel de activación: miedo no se diferencia significativamente del estado neutro en cuanto a la  $F_0$  media, variabilidad de la  $F_0$ , energía media, variabilidad de la energía media y duración, por lo que es considerada como una emoción pasiva; cólera puntúa, sin embargo, por encima del estado neutro en todos estos parámetros excepto en la  $F_0$  media donde la situación es comparable con la de miedo.

La confusión cólera-miedo aparece también en algunos estudios anteriores (Davitz y Davitz, 1964; Kramer, 1964) y es explicada en términos de nivel de activación. Esta discrepancia no es tan sorprendente si consideramos que la propia caracterización acústica de miedo varía de unos trabajos a otros. Los datos descriptivos de las expresiones utilizadas por nosotros la señalan como una emoción caracterizada por bajos niveles de actividad. Existen, sin embargo, otros estudios que informan de valores altos de tono y rango de tono para las expresiones de esta emoción (Fairbanks y Pronovost, 1939; Williams y Stevens, 1972), valores que coinciden con las predicciones teóricas de Scherer (1984) a partir de su modelo de procesos componentes.

Estas divergencias podrían ser explicadas con cargo a dicho modelo, ya que es posible que bajo la etiqueta «miedo» se incluyan dos estados emocio-

nales que podrían emerger de los distintos resultados ante uno de los chequeos realizados por el subsistema de procesamiento de la información, el relativo a la posibilidad de control de la situación (Jiménez Fernández, 1985).

## EXPERIMENTO II

### Método

Los sujetos utilizados como jueces fueron los mismos que realizaron la tarea de reconocimiento de las frases sin modificar, permaneciendo también idénticos el diseño y el procedimiento, por lo cual no van a ser descritos de nuevo en este apartado.

Las expresiones modificadas constituían también un total de 24, ya que las seis frases neutras fueron aproximadas en los parámetros de frecuencia fundamental a los valores medios que caracterizaban a las expresiones originales, de tal forma que para cada estado emocional se presentaron seis frases modificadas.

Los efectos de la posición serial de los estímulos fueron controlados del modo ya descrito para las expresiones sin modificar, obteniéndose las siguientes secuencias:

Bloque 1: N2M, N5A, N1T, N6A, N1C, N5T, N4C, N3T  
 Bloque 2: N3C, N3A, N6M, N2A, N5M, N6T, N5C, N6C  
 Bloque 3: N1M, N2T, N1A, N4M, N4A, N4T, N3M, N2C

Los dos primeros dígitos indican de qué frases neutras se trata, mientras que la última letra representa la inicial del perfil emocional al que se intentó aproximar.

### Resultados

La cuantificación de las respuestas se realizó siguiendo el modelo ya descrito para las expresiones sin modificar, obteniéndose las matrices de datos (ver tabla VI) en valores absolutos y proporciones. Se adjuntan también los resultados procedentes de los experimentos que suscitaron la réplica (ver tabla VII), a fin de disponer de una base que permita la comparación de los resultados y las estimaciones de fiabilidad de los juicios (ver tabla VIII).

TABLA VI

*Matriz completa de datos para todos los jueces empleando frecuencias absolutas (tabla superior) y frecuencias relativas expresadas en proporciones (tabla inferior)*

Categoría elegida	Categoría correcta				Total
	Alegría	Tristeza	Miedo	Cólera	
Alegría	102	30	107	22	261
Tristeza	76	264	104	167	611
Miedo	103	116	147	150	516
Cólera	248	123	167	196	734
No contesta	11	7	15	5	38
Total	540	540	540	540	2.160

Categoría elegida	Categoría correcta				Total
	Alegría	Tristeza	Miedo	Cólera	
Alegría	0.189	0.055	0.198	0.041	0.483
Tristeza	0.140	0.489	0.192	0.309	1.131
Miedo	0.191	0.215	0.272	0.278	0.955
Cólera	0.459	0.228	0.309	0.363	1.359
No contesta	0.020	0.013	0.028	0.009	0.070
Total	1.000	1.000	1.000	1.000	4.000

En este experimento, nuestro interés radica en comprobar si los sujetos son capaces de reconocer como emocionales frases originalmente neutras, cuyos perfiles de tono fueron artificialmente aproximados a los que caracterizaron las distintas emociones. Los análisis inferenciales se realizaron de nuevo sobre la media de frase reconocidas, media que se contrastó con la esperada por azar (1.5).

La media total de frases identificadas correctamente, considerando conjuntamente los cuatro tratamientos, resulta significativamente superior a la esperada por azar ( $t(89) = 3.19$ ,  $p < .01$ ) ello no implica, sin embargo, que todas las emociones hayan sido reconocidas. De hecho, las frases aproximadas a alegría no fueron identificadas como tales ( $t(89) = -2.60$ ,  $p > .05$ ). Tristeza y cólera, por el contrario, mostraron una media de reconocimiento altamente significativa ( $t(89) = 9.46$ ,  $p < .01$  y  $t(89) = 5.29$ ,  $p < .01$ , respectivamente). En el caso de miedo, cabe señalar que la puntuación obtenida es superior a la del azar, aunque la diferencia entre ambas no tiene la magnitud suficiente para alcanzar el nivel de significación estadística (ver tabla IX).

TABLA VII

*Matriz completa de datos (Jiménez Fernández, 1985) para todos los jueces empleando frecuencias absolutas (tabla superior) y frecuencias relativas expresadas en proporciones (tabla inferior)*

Categoría elegida	Categoría correcta				Total
	Alegría	Tristeza	Miedo	Cólera	
Alegría	35	8	40	11	94
Tristeza	37	63	32	43	175
Miedo	69	36	64	66	235
Cólera	92	59	40	106	297
Neutra	271	346	342	277	1236
No reconoce	30	22	16	31	99
Total	534	534	534	534	2136

Categoría elegida	Categoría correcta				Total
	Alegría	Tristeza	Miedo	Cólera	
Alegría	0.065	0.015	0.075	0.021	0.176
Tristeza	0.069	0.118	0.060	0.080	0.328
Miedo	0.129	0.067	0.120	0.124	0.440
Cólera	0.172	0.110	0.075	0.198	0.556
Neutra	0.507	0.648	0.640	0.519	2.315
No reconoce	0.056	0.041	0.030	0.058	0.185
Total	1.000	1.000	1.000	1.000	4.000

TABLA VIII

*Estimación de los índices de fiabilidad (efectivo: R y fiabilidad media: r) entre los 90 jueces en sus valoraciones de reconocimiento de las frases modificadas (por emociones y para el total del experimento)*

Fuente	SC	GL	MS	r (est.)	R (est.)
Juicios	4.60000	5	0.92000	0.07134	0.87365
Jueces	26.40000	89	0.29663		
Residuo	51.73333	445	0.11625	Alegría	
Juicios	3.02222	5	0.60444	0.01805	0.62337
Jueces	30.60000	89	0.34382		
Residuo	101.31111	445	0.22767	Tristeza	
Juicios	13.52778	5	2.70556	0.15391	0.94243
Jueces	24.15000	89	0.27135		
Residuo	69.30556	445	0.15574	Miedo	
Juicios	33.79259	5	6.75852	0.32053	0.97698
Jueces	21.85962	89	0.24610		
Residuo	69.20741	445	0.15552	Cólera	
Juicios	81.71065	23	3.55264	0.1695	0.94838
Jueces	19.23565	89	0.21613		
Residuo	375.33102	2047	0.18336	Total	experimento

TABLA IX

*Media de frases modificadas reconocidas y desviaciones típicas. El asterisco indica que valores resultaron estadísticamente significativos ( $p < .01$ )*

Emoción	Media	Desviación típica
Alegría	1.133	1.334
Tristeza	2.933*	1.436
Miedo	1.633	1.276
Cólera	2.178*	1.214
Total	1.969	1473

Estos datos son muy interesantes, especialmente al compararlos con los obtenidos por Jiménez Fernández (1985) para estas mismas frases (ver figura 3). En su trabajo ninguna de las emociones fue reconocida por encima del azar; alegría, tristeza y miedo presentaron puntuaciones inferiores a la esperada por azar, cólera obtuvo una puntuación ligeramente superior, pero que no resultó significativa.

Un segundo ANOVA fue llevado a cabo con el fin de averiguar si existen diferencias significativas entre los cuatro «estados emocionales» respecto a la media de frases modificadas que son identificadas correctamente. El hecho mismo de que unas emociones resultaran reconocidas por encima del azar, mientras que otras no, hacía ya sospechar de tales diferencias. Los resultados del ANOVA corroboraron la sospecha, mostrando un efecto significativo del factor emoción ( $F(89) = 28.48, p < .001$ ).

Del test de comparaciones múltiples se desprende que tristeza es significativamente mejor reconocida que el resto de las emociones (49 por 100 de elecciones correctas, 2.93 como media de frases identificadas). No se encon-



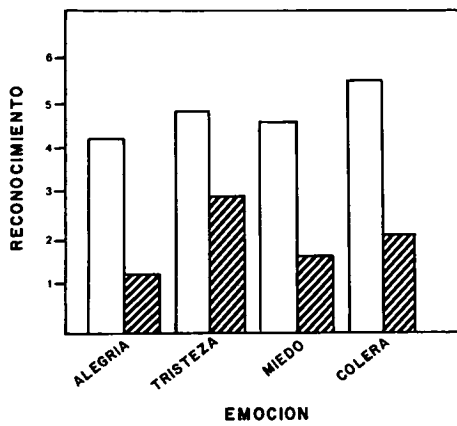


FIGURA 3

*Comparación de las puntuaciones medias de reconocimiento obtenidas en este trabajo para las frases modificadas con las obtenidas por Jiménez Fernández, 1985 (área rayada).*

traron, sin embargo, diferencias entre cólera (36 por 100, 2.17) y miedo (27 por 100, 1.63); ni entre esta última y alegría (19 por 100, 1.13).

Los resultados de la corrección por tendencia de respuesta pueden observarse en la tercera columna de la tabla X. Los rangos de los porcentajes de reconocimiento no se ven excesivamente alterados tras la corrección, únicamente miedo y alegría permutan sus posiciones, siendo estos rangos bastante concordantes con los del experimento replicado, con excepción de que en aquel trabajo, cólera fue la emoción que obtuvo un mayor número de elecciones correctas, seguida de tristeza (ver tabla XI). Recordemos que a pesar de esta gradación, ninguna de las dos emociones superó la barrera del azar.

Finalmente, quisimos comparar las frases originales y modificadas respecto a su grado de reconocimiento; dado que ambas habían sido juzgadas por los mismos sujetos. En la figura 4 se representan mediante un diagrama de

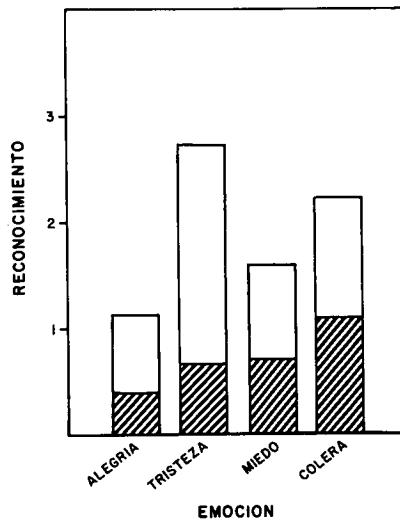


FIGURA 4

*Comparación de las puntuaciones medias de reconocimiento entre las frases originales (barra clara) y las modificadas (barra rayada).*

TABLA X

*Porcentaje de elecciones correctas para cada categoría emocional, incluyendo los resultados tras la corrección por azar, la corrección por tendencia de respuesta y el porcentaje total de elecciones. El número entre paréntesis indica el rango, desde 1 (emoción mejor reconocida) hasta 4 (emoción peor reconocida)*

Emoción	Elecciones correctas (%)	Total de elecciones (%)	Corrección tendencia de respuesta (%)
Alegría	19 (4)	48	29 (3)
Tristeza	49 (2)	113	46 (1)
Miedo	27 (3)	95	28 (4)
Cólera	36 (2)	136	31 (2)

TABLA XI

*Porcentaje de elecciones correctas obtenido por Jiménez Fernández (1985) incluyendo los resultados tras la corrección por azar, la corrección por tendencia de respuesta y el porcentaje total de elecciones. El número entre paréntesis indica el rango, desde 1 (emoción mejor reconocida) hasta 4 (emoción peor reconocida)*

Emoción	Elecciones correctas (%)	Total de elecciones (%)	Corrección tendencia de respuesta (%)
Alegría	6 (4)	17	22 (3)
Tristeza	11 (3)	33	24 (2)
Miedo	12 (2)	44	19 (4)
Cólera	20 (1)	55	28 (1)

barras las puntuaciones obtenidas por cada emoción con los dos tipos de frases. Se realizó un ANOVA de dos factores ( $4 \times 2$ ) siendo el primer factor el tipo de emoción y el segundo la condición de modificación frente a no modificación. Los resultados indican que existen diferencias altamente significativas en el reconocimiento debidas tanto al tipo de emoción considerada ( $F(3)=41.7$ ,  $p<.0001$ ) como a la variable modificación ( $F(1)=1020.87$ ,  $p<.0001$ ). Asimismo, existe efecto de interacción ( $F(3)=15.01$ ,  $p<.0001$ ) lo que indica que la condición modificación afecta diferencialmente a las distintas emociones, dato que se desprendía ya de los resultados expuestos en los epígrafes anteriores.

## Discusión

El objetivo principal de este trabajo consistía en averiguar si los sujetos eran capaces de identificar el significado emocional de las expresiones vocales basándose exclusivamente en la información transportada a través de las señales del perfil de tono. Dado que las frases neutras mantenían sus valores característicos en el resto de los parámetros, su reconocimiento como «emocionales» por parte de los sujetos indicaría la importancia del perfil de tono en la comprensión del contenido afectivo del mensaje.

Nuestros resultados muestran que cuando los sujetos se ven forzados a elegir entre las categorías emocionales por no disponer de las opciones «Neutra» y «No reconoce», se produce una mejoría en el reconocimiento respecto

al experimento anterior en el cual estas opciones estaban incluidas. En nuestro trabajo, dos emociones, tristeza con un 49 por 100 de elecciones correctas, y cólera con un 36 por 100, fueron reconocidas por encima del azar; miedo con un 27 por 100 de aciertos se situó también por encima del azar, pero no de forma significativa. Esto supone una confirmación parcial de la hipótesis que resulta de difícil interpretación. Una posibilidad consistiría en suponer que si tristeza y cólera han sido identificadas, es porque las señales vocales manipuladas son más relevantes para el reconocimiento de estas emociones que para la identificación de alegría y miedo.

Esta idea es difícil de confirmar dado el estado actual de la investigación, donde no existe acuerdo entre los distintos estudios respecto a qué señales acústicas específicas conducen a los decodificadores a la atribución de estados afectivos discretos. Las únicas conclusiones bien establecidas se refieren a la relación de algunos parámetros con dimensiones más generales; en el trabajo de Ladd, Silverman, Tolkmitt, Bergmann y Scherer (1985) se señala el papel del rango de tono en las atribuciones de activación, mientras que los parámetros relacionados con la cualidad de la voz tendrían una mayor intervención en los juicios sobre la dimensión evaluativa.

A modo de sugerencia planteamos que es precisamente en las emociones reconocidas en nuestro experimento donde existe un mayor número de trabajos que concuerdan respecto a cómo se comportan a nivel expresivo las variables relacionadas con el tono: nivel de tono y rango. En la revisión realizada por Scherer (1981), siete estudios (Fairbanks y Pronovost, 1939; Eldred y Price, 1958; Sedláček y Sychra, 1963; Davitz, 1964a; Huttar, 1968; Williams y Stevens, 1972 y Scherer y Walbott, en prensa) señalan que tristeza se caracteriza por un bajo nivel de tono, y seis trabajos (Fairbanks y Pronovost; Eldred y Price; Davitz; Huttar; Williams y Stevens; Levin y Lord, 1975) encontraron un alto nivel de tono para cólera. El incremento de tono para las emisiones de alegría y miedo únicamente es referido por cuatro y dos estudios respectivamente. En cuanto al rango de tono, existen menos datos disponibles para todas las emociones, pero al menos los trabajos de Fairbanks y Pronovost, y William y Stevens, coinciden en indicar un rango de tono reducido para tristeza y amplio para cólera, mientras que no existe ninguna información en relación a este parámetro para alegría y únicamente un estudio señala que miedo se caracteriza por un rango amplio de tono.

Si los parámetros de tono tuvieran un mayor peso para discriminar las emociones de tristeza y cólera es lógico que, al modificar las frases neutras únicamente en tales parámetros, dichas emociones se reconozcan mejor que otras, como miedo y alegría cuya identificación podría descansar en otro tipo de variables.

En relación con esta idea recordemos las sugerencias de Van Bezooijen (1984) con respecto a la emoción de alegría; esta emoción que en términos perceptivos era juzgada de forma similar en varias lenguas, obtuvo, sin embargo, puntuaciones de reconocimiento radicalmente diferentes en las tres culturas estudiadas: fue muy bien identificada cuando la lengua del hablante y del oyente coincidían y una de las peor reconocidas en caso contrario. El autor señala que es posible que esta emoción, además de por los cuatro parámetros considerados al hacer la comparación en términos perceptivos (tono, rango de tono, sonoridad y ritmo), esté caracterizada por señales más sutiles que varíen de cultura a cultura, señales que según el autor podrían ser de naturaleza espectral. Esta interpretación concuerda con el hecho de que en nuestro experimento alegría no haya sido reconocida cuando la única pista ofrecida a los sujetos para su identificación fue su perfil de tono.

A pesar de estas consideraciones es necesario señalar que los datos obte-

nidos deben ser interpretados con cautela, ya que las codificaciones fueron realizadas por un único sujeto y los resultados podrían no ser lo suficientemente generales. Cabe la posibilidad de que las características particulares de la voz del codificador indujeran a los jueces a atribuciones de determinados estados emocionales en detrimento de otros.

Por otra parte, las frases neutras fueron aproximadas en su perfil de tono a los valores medios de las expresiones juzgadas como emocionales. Probablemente hubiera sido deseable aproximarlas a los valores de la frase mejor reconocida de cada categoría emocional dado que una frase mal codificada podría distorsionar el valor del perfil medio para esa emoción. De hecho, se encontraron diferencias significativas entre las seis frases originales, respecto a su grado de reconocimiento, para las cuatro categorías emocionales.

## CONCLUSIONES

El trabajo realizado se integra en la línea de investigación interesada en el reconocimiento de emociones a través de las señales transportadas por el canal vocal. El experimento, que se llevó a cabo utilizando como material estimular las frases sin modificar, tiene como objetivo determinar si tal reconocimiento se produce y en qué grado. A partir de nuestros resultados podemos responder afirmativamente a la primera parte de esta cuestión, dado que todas las emociones fueron reconocidas por encima del azar, obteniéndose, además, porcentajes de aciertos que se sitúan entre los más elevados dentro de los ofrecidos por la literatura.

Se examinaron también algunos de los problemas que más han preocupado a la investigación en este campo: las diferencias en el grado de reconocimiento de las distintas emociones y las confusiones que se producen entre ellas. Respecto al primer tema, nuestros resultados indican que algunas emociones son significativamente mejor reconocidas que otras, obteniendo cólera los mayores porcentajes de identificación correcta. Por otra parte, las posiciones que ocupan las emociones estudiadas cuando son ordenadas en función del número de aciertos, se corresponden, en general, con las referidas por trabajos anteriores.

Explicar la razón de tales diferencias, resulta, sin embargo, bastante complicado y es una tarea que muy pocos investigadores han abordado. La única hipótesis al respecto propone un mejor reconocimiento de las emociones negativas frente a las positivas en términos del mayor «valor adaptativo» de aquéllas y aun así no todos los datos empíricos permiten sostener un peor reconocimiento para las emociones positivas, al menos por lo que respecta a la investigación en adultos. Los resultados de nuestro trabajo apoyarían en principio la hipótesis propuesta, aunque tal apoyo debe ser matizado en función de las consideraciones ya expuestas.

En cuanto al análisis de errores, pese a que en la literatura revisada se propone la dimensión actividad como la principal responsable de las confusiones entre emociones, los errores aquí encontrados se explican mejor en términos de similaridad en la dimensión evaluativa.

La utilización de frases neutras modificadas en su perfil de tono responde al intento de evaluar la contribución de parámetros acústicos concretos al reconocimiento de emociones. El resultado más interesante de este trabajo es, probablemente, que dos de ellas: tristeza y cólera, hayan sido reconocidas por encima del azar. Sin embargo, es difícil —a la luz de los actuales conocimientos— establecer explicaciones fundamentadas para este hallazgo, siendo esta ausencia de bases explicativas uno de los principales problemas con que se

ha enfrentado la mayor parte de la investigación en el área. Scherer (1985, 1986), tras hacer una breve revisión del estado actual de la investigación, señala la necesidad urgente de que sean razones teóricas las que motiven la elección de los parámetros vocales a estudiar en lugar de proceder como hasta ahora de una manera casi aleatoria. En su opinión esto sólo es posible si se profundiza en el conocimiento tanto de los patrones fisiológicos que caracterizan a los estados emocionales discretos, como de la neuropsicología de las expresiones vocales con significado afectivo. Insiste también en la necesidad de investigación en ciertas áreas (el enfoque comparativo y la investigación evolutiva), hasta ahora abandonadas, que pueden proporcionar importantes claves teóricas para la comprensión de los fenómenos que nos ocupan.

## Referencias

- ALBAS, D. C.; MCCLUSKEY, L. W., y ALBAS, C. A.: «Perception of the emotional content of speech: A comparison of two Canadian groups». *Journal of Cross-Cultural Psychology*, 1976, 7, 481-490.
- BLANCK, P. D.; ROSENTHAL, R.; SNODGRASS, S. E.; DEPAULO, B. M., y ZUCKERMAN, M.: «Sex differences in eavesdropping on nonverbal cues: Developmental changes». *Journal of Personality and Social Psychology*, 1981, 41, 391-396.
- BEIER, E. C., y ZAUTRA, A. J.: «Identification of vocal communication of emotions across cultures». *Journal of Consulting and Clinical Psychology*, 1972, 39, 166.
- BEZOOIJEN, R. VAN: *Characteristics and Recognizability of vocal expression of emotion*. Dordrecht (Holland): Foris Publications, 1984.
- BRIGHETTI, G.; LADAVAS, E., y RICCI BITTI, P. E.: «Recognition of emotion expressed through voice». *Giornale Italiano di Psicologia*, 1980, 7, 121-127.
- BROWN, B. L., y BRADSHAW, J. M.: «Towards a Social Psychology of Voice Variations», en H. Giles y R. N. St. Clair (Eds.): *Recent advances in Language, Communication and Social Psychology*. Londres: Erlbaum, 1985.
- BURNS, K. L., y BEIER, E. G.: «Significance of vocal and visual channels in the decoding of emotional meaning». *Journal of Communication*, 1973, 23, 118-130.
- DAVITZ, J. R., y DAVITZ, L. J.: «The communications of feelings by content-free speech». *Journal of Communication*, 1959, 9, 6-13.
- DAVITZ, J. R. (Ed.): *The communication of emotional meaning*, Nueva York: McGraw-Hill, 1964.
- DIMITROSKY, L.: «The ability to identify the emotional meaning of vocal expressions at successive ages», en J. R. Davitz (Ed.): *The communication of emotional meaning*. Nueva York: McGraw-Hill, 1964.
- DUSENBURY, D., y KNOWER, F. H.: «Experimental studies of the symbolism of action and voice (III). A study of the specificity of meaning in facial expression». *Quarterly Journal of Speech*, 1938, 24, 424-435.
- FAIRBANKS, G.: «Recent experimental investigations of vocal pitch in voice». *Journal of Acoustical Society of America*, 1940, 11, 457-466.
- FAIRBANKS, G., y HOAGLIN, L. W.: «An experimental study of the durational characteristics of the voice during the expression of emotion». *Speech Monographs*, 1941, 8, 85-90.
- FAIRBANKS, G., y PRONOVOST, W.: «An experimental study of the pitch characteristics of the voice during the expression of emotion». *Speech Monographs*, 1939, 6, 87-104.
- FECHNER, E. H.: «Children understanding of the nonverbal communication of emotions in the visual, vocal and gestural». *Dissertation Abstracts International*, 1978, 39 (5-9), 2568-2569.
- FELDSTEIN, S.; JAFFE, J., y CASSOTTA, L.: «The discrimination and identification of simulated affective vocal behavior». *Research Bulletin*, 1968, 3.
- FENSTER, C. A.: «Vocal communication of emotional meaning among adults and children». *Dissertation Abstracts International*, 1967, 28 (4-B), 1694-1695.
- FENSTER, C. A.; BLAKE, L. K., y GOLDSTEIN, A. M.: «Accuracy of vocal emotional communications among children and adults and the power of negative emotions». *Journal of Communication Disorders*, 1977, 10, 301-314.
- HAUSER, K. O.: «The use of acoustical analysis for identification of client stress within the counseling session». *Dissertation Abstracts International*, 1976, 36, 5149-5150.
- JIMÉNEZ FERNÁNDEZ, A.: *Marcadores emocionales en la conducta vocal*. Tesis doctoral no publicada. Universidad Autónoma de Madrid, 1985.
- KRAMER, E.: «Judgments of personal characteristics and emotions from nonverbal properties of speech». *Psychological Bulletin*, 1963, 60, 408-420.
- KRAMER, E.: «Elimination of nonverbal cues in judgements of emotion from voice». *Journal of Abnormal Social Psychology*, 1964, 68, 390-396.
- LADD, R.; SILVERMAN, K. E. A.; TOLKMITT, F.; BERGMANN, G., y SCHERER, K. R.: «Evidence for the independent function of intonation contour type, voice quality and Fo range in signalling speaker affect». *Journal of the Acoustical Society of America*, 1985, 78, 435-444.
- LEVIN, H., y LORD, W.: «Speech pith frequency as an emotional state indicator». *IEEE Transactions and Systems, Man and Cybernetics*, SMC-5, 1975, 2, 259-272.
- LEWITT, E. A.: «The relationships between abilities to express emotional meaning vocally and facially», en J. R. Davitz (Ed.): *The communications of emotional meaning*. Nueva York: McGraw-Hill, 1964.
- MALLARD, A. R., y DANILOFF, R. G.: «Glottal cues for parents judgment of emotional aspects of infant vocalizations». *Journal of Speech and Hearing Research*, 1979, 16, 592-596.

- MCCCLUSKEY, K. W.; ALBAS, D. C., y NIEMI, R. R.: «Cross cultural differences in the perception of the emotional content of speech: A study of the development of sensitivity in Canadian and Mexican Children». *Developmental Psychology*, 1978, 11, 551-555.
- MUÑOZ, C.: *Reconocimiento de patrones en la expresión emocional a través de la conducta vocal*. Memoria de Licenciatura no publicada. Universidad Autónoma de Madrid, 1986.
- PEAFF, P. L.: «An experimental study of the communications of feelings without contextual materials». *Speech Monographs*, 1954, 21, 155-156.
- ROSENTHAL, R.: «Conducting judgments studies», en K. R. Scherer y P. Ekman (Eds.): *Handbook of Methods in Nonverbal Behavior Research*. Cambridge: Cambridge University Press, 1982.
- ROSS, M.; DUFFY, R. J.; COOKER, H. S., y SARGEAN, R. L.: «Contribution of the lower audible frequencies to the recognition of emotions». *American Annals of the Deaf*, 1973, 118, 37-62.
- SCHERER, K. R.: «Nonlinguistic vocal indicators of emotion and psychopathology», en C. E. Izard (Ed.): *Emotions in Personality and Psychopathology*. Nueva York: Plenum, 1979.
- SCHERER, K. R.: «Speech and Emotional States», en J. K. Darby (Ed.): *Speech evaluation in Psychiatry*, Nueva York: Grune Stratton, 1981.
- SCHERER, K. R.: «Methods of research on vocal communication: paradigms and parameters», en K. R. Scherer y P. Ekman (Eds.): *Handbook of methods in nonverbal behavior research* Cambridge: Cambridge University Press, 1982.
- SCHERER, K. R.: «On the nature and function of emotion: A Component Process Approach». En K. R. Scherer y P. Ekman: *Approaches to emotion*. Hillsdale (N. J.): Erlbaum, 1984.
- SCHERER, K. R.: «The vocal expression of emotion: A progress report». *Revista de Psicología Social*, 1985, 0, 51-54.
- SCHERER, K. R.: «Vocal affect expression: a review and a model for future research». *Psychological Bulletin*, 1986, 99, 143-165.
- SCHERER, K. R.; KOIVUMAKI, J., y ROSENTHAL, R.: «Minimal Cues in the vocal Communication of affect: Judging emotions from content-masked speech». *Journal of Psycholinguistics Research*, 1972, 1, 259-265.
- SCHERER, K. R.; LADD, D. R., y SILVERMAN, K. E. A.: «Vocal cues to speaker affect: Testing two models». *Journal of the Acoustical Society of America*, 1984, 76, 1346-1356.
- SCHERER, K. R.; LONDON, H., y WOLF, J. J.: «The voice of confidence: Paralinguistic cues and Audience Evaluation». *Journal of Research in Personality*, 1975, 71, 31-44.
- SCHERER, K. R., y OSHINSKY: «Cue utilization in emotion attribution from auditory stimuli». *Motivation and Emotion*, 1977, 1, 331-346.
- SIMONOV, P. V., y FROLOV, M. B.: «Utilization of Human Voice for Estimation of Man's Emotional stress and state of attention». *Aerospace Medicine*, 1973, 44, 256-258.
- SIMONOV, P. V., y FROLOV, M. V.: «Analysis of the human voice as a method of controlling emotional state: Achievements and goals». *Aviation, Space, and Environmental Medicine*, 1977, 48, 23-25.
- SKINNER, E. R.: «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*, 1935, 2, 81-137.
- SOSKIN, W. F., y KAUFFMAN, P. E.: «Judgment of emotion in word-free voice samples». *The Journal of Communication*, 1961, 11, 73-80.
- STREETER, L. A.; KRAUSS, R. M.; APPLE, W., y MACDONALD, N. H.: «Acoustic consequences and perceptual indicators of stress». *Journal of the Acoustical Society of America*, 1978, 64, S115.
- WILLIAMS, C. E., y STEVENS, K. W.: «On determining the emotional state of pilots during flight: An exploratory study». *Aerospace Medicine*, 1969, 40, 1369-1372.
- WILLIAMS, C. E., y STEVENS, K. W.: «Emotions and speech. Some acoustical correlates». *The Journal of the Acoustical Society of America*, 1972, 52, 1238-1250.
- WILLIAMS, C. E., y STEVENS, K. W.: «Vocal correlates of emotional states», en J. K. Darby (Ed.): *Speech evaluation in psychiatry*. Nueva York: Grune & Stratton, 1981.
- WOOLF, G.; GORSKI, R., y PETERS, S.: «Acquaintance and accuracy of vocal communication of emotions». *Journal of Communication*, 1972, 22, 300-305.
- ZUCKERMAN, M.; LIPETS, M. S.; KOIVUMAKI, H. J., y ROSENTHAL, R.: «Encoding and decoding nonverbal cues of emotion». *Journal of Personality and Social Psychology*, 1975, 32, 1068-1076.