

Sesgos cognitivos en el reconocimiento de expresiones emocionales de voz sintética en la alexitimia

Francisco Martínez Sánchez, Juan Manuel Montero Martínez y Javier de la Cerra
Universidad de Murcia

En este trabajo investigamos la existencia de sesgos cognitivos en el reconocimiento de expresiones vocales emocionales, generadas usando procedimientos de conversión texto a voz por síntesis de formantes, en relación a la capacidad para procesar estímulos emocionales en la alexitimia, con el objeto de valorar si ésta refleja un déficit en el procesamiento emocional. Sometimos a un grupo formado por 162 sujetos a diversos estímulos en los que debían identificar la emoción (neutra, alegre, triste, airada) que expresaba el hablante. El análisis de las matrices de confusión de los resultados mostraron que los sujetos con altos niveles de alexitimia reconocían significativamente menos expresiones emocionales que quienes tenían bajos niveles de alexitimia durante la primera mitad del experimento. Estos resultados sugieren que los alexitímicos son menos hábiles para reconocer estímulos vocales emocionales, además proveen de validez al constructo de alexitimia.

Cognitive biases in the recognition of synthetic emotional speech in alexithymia. In this article, we research the cognitive biases in the recognition of synthetic emotional speech generated using formant-based text-to-speech conversion, related to emotion processing capacities in alexithymia, in order to assess that alexithymia reflects a deficit in the ability to process emotional stimuli. The different recordings of the stimuli were presented to the each listener (162 undergraduate students), and he/she had to identify the emotion. It was possible to choose among neutral, happy, sad, angry and unidentified. Operating on these results, we have obtained the confusion matrices of the identified emotions. Results show that the subjects categorized as alexithymics recognize significantly less emotional expressions than the subjects with low levels of alexithymia in the first half of the experiment. These findings suggest that alexithymics were less able to recognize emotional stimuli. Also these findings provide further support for the validity of the alexithymia construct.

La alexitimia –etimológicamente, ausencia de palabras para expresar las propias emociones– es un constructo hipotético multidimensional, formulado en la década de los setenta por Sifneos (1973) para describir una compleja constelación de manifestaciones cognitivo-afectivas observadas en pacientes aquejados de alteraciones psicósomáticas. El constructo ha suscitado gran interés entre los investigadores de la Psicología de la Emoción, así como para los interesados por la psicósomática, la psicología de la salud y la psicopatología.

Se considera que quienes padecen altos niveles de alexitimia muestran una marcada dificultad para identificar sentimientos y diferenciarlos de las sensaciones fisiológicas que acompañan a la activación emocional, junto a dificultades para describir sentimientos a los otros, asimismo muestran constricción en los procesos simbólicos, expresada a través de una reducida capacidad de fantasía, así como un patrón de expectativas y atribuciones orientado a los acontecimientos y detalles externos (Martínez-Sánchez,

1999; Taylor, 2000). Estas características son interpretadas en términos de un desorden en la regulación afectiva, así como un rasgo de personalidad (Taylor, Bagby y Parker, 1991) muy estable en el tiempo (Martínez-Sánchez, Ato, Córcoles, Huedo y Selva, 1998) capaz de influir en la reactividad fisiológica emocional (Infrasca, 1997), propiciando la aparición de alteraciones relacionadas con la hiperactivación fisiológica (Lumley, Stettner y Wehmer, 1996; Lumley, Tomakowsky y Torosian, 1997).

Si bien han elaborado numerosas hipótesis etiológicas para explicar su origen, más recientemente, diversos autores (Martin y Pihl, 1985; Martínez-Sánchez y Fernández Castro, 1994; Berenbaum y Prince, 1994; Roedema y Simons, 1999; Suslow, 1998) han propuesto la consideración de la alexitimia como un trastorno específico del procesamiento de la información emocional, manifestado mediante una marcada dificultad para procesar información de carácter afectivo, que provoca que las emociones se mantengan parcialmente indiferenciadas y pobremente reguladas (Taylor, Bagby y Parker, 1997).

Recientemente han sido validadas experimentalmente algunas de las premisas sobre las que se asienta esta hipótesis, entre las que destacan: 1) Dificultad para procesar información afectiva de carácter no lingüístico (Bradshaw, 1989; Dewaraja y Sasaki, 1990), tales como expresiones faciales emocionales (McDonald y Prkachin, 1990; Mandal y Singh, 1990; Jessimer y Markham, 1997); 2) Difi-

cultad para discriminar entre distintos estados emocionales en tareas que implican el uso de descriptores verbales (Bagby, Parker, Taylor y Acklin, 1993; Lamberty y Holt, 1995); 3) Patrones atencionales específicos de la información emocional (Martínez Sánchez y Marín, 1997); 4) Procesamiento no simbólico de la información visual (Montreuil y Jouvent, 1989; Montreuil, Jouvent, Carton, Bungenier y Widlocher, 1991) y 5) Dificultades en la propiocepción visceral de las manifestaciones fisiológicas asociadas a la activación emocional (Martínez-Sánchez, Ortiz y Ato, 2001; Näring y van der Staak, 1995; Sachse, 1994); y 6) Patrones específicos de activación en respuesta a estímulos afectivos (Berembaum y Prince, 1994; Parker, Taylor y Bagby, 1992; Silberman y Weingartner, 1986).

Con el objeto de valorar la existencia de déficits en el procesamiento de estímulos afectivos en relación con la alexitimia, nos propusimos con este trabajo valorar la capacidad para identificar expresiones vocales emocionales en sujetos con diversos niveles de alexitimia. Extremo desconocido actualmente y que podría, a nuestro juicio, contribuir al apoyo de la denominada «hipótesis cognitiva». Hipotetizamos que las tasas de reconocimiento de los estímulos con prosodia emocional serán significativamente menores en el grupo con altos niveles de alexitimia. Para ello ideamos un procedimiento experimental consistente en una tarea típica de reconocimiento de expresiones vocales emocionales (Burkhardt y Sendlmeier, 2000), para lo que empleamos expresiones generadas por procedimientos de conversión texto-voz, por entrañar esta tarea mayor dificultad en la identificación por parte de los oyentes que la voz natural, efecto provocado por la propia naturaleza y el proceso de implementación de los estímulos de naturaleza sintética.

La conversión texto-voz (CTV), en inglés Text-To-Speech (TTS), es el proceso automático por el que se genera voz artificial –sintética– a partir de un texto, simulando así la conducta vocal de un lector o intérprete (Klatt, 1987). Diversos parámetros determinan la calidad de la voz sintética, entre los que destacamos su inteligibilidad (porcentaje de palabras correctamente identificadas), así como su naturalidad (grado de similitud con la voz humana natural). Si bien actualmente la inteligibilidad es un objetivo plenamente alcanzado en la mayoría de los sistemas desarrollados, en los que se obtienen tasas superiores al noventa y cinco por ciento, la naturalidad es aún materia de investigación (Montero, Gutiérrez-Arriola, Colás, Enríquez y Pardo, 1999), dadas las dificultades que plantea la implementación de sistemas que reproduzcan con naturalidad la voz sintética.

Uno de los principales factores que afectan a la naturalidad de la voz sintética es la monotonía y la prosodia ausente de matiz afectivo con la que es percibida por los oyentes (Murray y Arnott, 1995). Este hecho ha propiciado recientemente la investigación de procedimientos que doten a los CTV de la capacidad de simular o transmitir emociones a través de la voz (Montero, Gutiérrez-Arriola, de Córdoba, Enríquez y Pardo, 2001). Dos son las principales técnicas de síntesis de CTV, por formantes y por concatenación. La primera emplea un modelo matemático del proceso de producción de voz que calcula los parámetros que caracterizan cada sonido por medio de un conjunto de reglas que valoran el contexto fonético de emisión de cada sonido para, finalmente, generar voz a partir de estos parámetros. En la técnica de síntesis por concatenación se parte de voz humana previamente grabada, se seleccionan las secuencias de fragmentos de voz a concatenar para producir el discurso deseado, se adapta cada fragmento de voz al contexto prosódico en que se desea insertar y, finalmente, se concatenan (Rank y Pirker, 1998).

Aunque la calidad de los sistemas basados en la concatenación es actualmente superior, la síntesis por formantes (Murray y Arnott, 1996), debido a su naturaleza paramétrica, es más frecuentemente empleada en síntesis de voz emocional, ya que permite sintetizarla a partir de una muestra de voz emocionalmente neutra, mediante la adaptación de los parámetros prosódicos perceptuales típicos que caracterizan la voz asociada a cada emoción.

Si bien la calidad de la voz sintética es inferior a la humana grabada digitalmente, la CTV resulta extraordinariamente útil en aplicaciones en las que el conjunto de mensajes a emitir es muy grande o variable, ya que resulta muy costoso grabar todos los mensajes que necesitaría un sistema que ha de leer, por ejemplo, nombres de personas o ciudades, especialmente si tenemos en cuenta que la producción de cada palabra depende fuertemente del contexto en que se inserta, lo que obligaría a realizar numerosas grabaciones de cada expresión.

Método

Sujetos

La muestra estuvo formada por 162 sujetos (36 hombres y 126 mujeres), de edades comprendidas entre los 18 y 40 años (\bar{x} = 19.93; s_x = 3.51), todos ellos asistentes a la asignatura de Psicología de la Motivación y la Emoción de la Universidad de Murcia.

Materiales

Escala de Alexitimia de Toronto (TAS-20) (Bagby, Parker y Taylor, 1994), en la adaptación española de Martínez-Sánchez (1996). La TAS-20 es un autoinforme de veinte ítems ideado para evaluar el nivel de alexitimia; está compuesto por tres factores congruentes con el constructo: (F. I) dificultad para identificar sentimientos y distinguir entre sentimientos y los síntomas asociados a la activación emocional; (F. II) dificultad para describir sentimientos a los otros y (F. III) pensamiento orientado a lo externo. La adaptación española de la escala fue realizada por psicólogos que hablaban fluidamente español e inglés, empleando el procedimiento de traducción inversa. Esta adaptación posee una adecuada consistencia interna (α = .78), así como elevada fiabilidad test-retest (r = .71, p < .001) en un intervalo de 19 semanas, comparables a los que posee la versión inglesa de la escala original. De las diversas técnicas empleadas para la evaluación de la alexitimia, la TAS-20 es la más frecuentemente empleada por sus destacadas propiedades psicométricas (Bagby, Taylor y Parker 1994; Martínez-Sánchez y Ortiz, 2000). La estabilidad y replicabilidad de su estructura factorial ha sido confirmada tanto en muestras clínicas como no clínicas (Bagby, Parker y Taylor, 1994; Páez et al., 1999; Parker, Bagby, Taylor, Endler y Schmitz, 1993).

Base de datos de voz sintética de expresiones emocionales elaborada dentro del proyecto europeo TIDE TP 1174 VAESS (*Voices, Attitudes and Emotions in Speech Synthesis*), por el Departamento de Ingeniería Electrónica de la ETSI de Telecomunicaciones de la Universidad Politécnica de Madrid (Montero, Gutiérrez-Arriola, Palazuelos-Cagigas, Enríquez, Aguilera y Pardo, 1998). La base de expresiones emocionales está compuesta por expresiones de voz sintética masculina, obtenidas mediante conversión automática de texto a voz. Los textos de las muestras carecen de sentido emocional. Para la generación de la voz sintética se ha partido de un sintetizador por formantes de voz neutra desarrollado

dentro del proyecto VAESS, a la que se han añadido reglas que, modificando la prosodia (intensidad, tono medio, velocidad, rango del tono, etc.) y la fuente glotal de voz neutra, simulan voz emitida en los estados emocionales de ira, tristeza y alegría, así como un estado emocional neutro.

Los estímulos auditivos empleados fueron dos frases: 1ª «Tengo la llave en el bolsillo» (1,573 seg.) y 2ª «Los participantes en el congreso marcharon después al Escorial» (3,869 seg.), cada una de ellas con muestras de voz que representan ira, tristeza, alegría, así como voz emocionalmente neutra. Los estímulos (muestreados a 16 KHz, con 16 bits de resolución PCM y formato monoaural) fueron presentados empleando unos auriculares Sennheiser modelo HD 495, desde un Minidisc Sharp modelo MD-R1HM MK II.

Procedimiento

Los resultados obtenidos tras la aplicación de la Escala de Alexitimia de Toronto permitieron clasificar a los sujetos en tres grupos atendiendo al criterio de los cuartiles. Quienes obtuvieron puntuación inferiores al primer cuartil ($Q_1 = 38$) se agregaron al grupo de «baja alexitimia» (11 hombres y 25 mujeres), mientras que quienes obtuvieron puntuaciones superiores al tercer cuartil ($Q_3 = 51$) conformaron el grupo de «alta alexitimia» (7 hombres y 30 mujeres). El resto de los sujetos, cuyas puntuaciones estuvieron comprendidas entre ambos puntos de corte, formaron el grupo de «alexitimia media» (18 hombres y 30 mujeres). La distribución de las puntuaciones obtenidas en la Escala de Alexitimia de Toronto se encuentran dentro del rango de valores obtenido en la adaptación de la escala (Martínez-Sánchez, 1996).

El procedimiento experimental consistió en la aplicación individual, bajo condiciones controladas de laboratorio, de ocho estímulos previamente aleatorizados, cuatro para la frase: «Tengo la llave en el bolsillo» e igual número para «Los participantes en el congreso marcharon después al Escorial»; se pedía a los sujetos que identificaran y registraran, en un protocolo confeccionado para tal fin, la emoción que expresaba el hablante: neutra, ira, tristeza o alegría; los sujetos tenían asimismo la opción de responder con la opción «no sé» en caso de no identificarla; en ningún caso se ofreció la posibilidad de repetir ningún estímulo. Previa a la fase experimental se realizó una fase de prueba, compuesta por cuatro estímulos de las mismas características que los que posteriormente se presentarían, con el objeto de familiarizarlos con la tarea y la fuente de voz sintética, en la que se presentó la frase «No queda fruta los viernes» (1,529 seg.).

Resultados

Inicialmente analizamos las tasas de reconocimiento de las dos frases empleadas. Como puede apreciarse en las matrices de confusión de las tablas 1 y 2, las mayores tasas de reconocimiento se dan en la expresión neutra y en la de tristeza, mientras que las menores corresponden a la alegría y la ira. Es preciso señalar la diferencia observada en las tasas de reconocimiento en ambas frases; así, las tasas se incrementan sensiblemente en la frase: «Los participantes en el congreso marcharon después al Escorial» (3,869 seg.), frente a las obtenidas en la frase: «Tengo la llave en el bolsillo» (1,573 seg.). Atribuimos este efecto a la diferente duración de ambas, por lo que en la frase de mayor duración el sujeto dispone de un mayor tiempo para identificar la prosodia emocional con la que está compuesta.

Tabla 1				
Matriz de confusión de las tasas de reconocimiento de la frase: «Tengo la llave en el bolsillo»				
Emoción identificada	Emoción simulada			
	Neutra	Tristeza	Alegría	Ira
Neutra	71,6% (116)	39,5% (64)	38,9% (63)	35,2% (57)
Tristeza	14,8% (24)	48,8% (79)		1,9% (3)
Alegría	,6% (1)	1,2% (2)	42,6% (69)	9,3% (15)
Ira	5,6% (9)	1,2% (2)	5,6% (9)	35,8% (58)
«No sé»	7,4% (12)	9,3% (15)	13,0% (21)	17,9% (29)

Nota: entre paréntesis aparece el número absoluto de reconocimientos

Tabla 2				
Matriz de confusión de las tasas de reconocimiento de la frase: «Los participantes en el congreso marcharon después al Escorial»				
Emoción identificada	Emoción simulada			
	Neutra	Tristeza	Alegría	Ira
Neutra	90,7% (147)	39,5% (64)	32,6% (53)	39,5% (64)
Tristeza		50,0% (81)		
Alegría			37,7% (61)	12,3% (20)
Ira	2,5% (4)	3,1% (5)	9,3% (15)	41,4% (67)
«No sé»	6,8% (11)	7,4% (12)	20,4% (33)	6,8% (11)

Nota: entre paréntesis aparece el número absoluto de reconocimientos

Se sometió a prueba empírica, empleando pruebas de comparación de proporciones con corrección de continuidad, si las proporciones de la diagonal principal de las ambas tablas de contingencia eran significativamente diferentes de las proporciones correspondientes a la categoría «neutra» para las emociones identificadas. En ambas tablas de contingencia observamos la misma pauta de resultados; así, para la categoría «tristeza», las proporciones diferían significativamente ($p < .05$), mientras que para las restantes categorías («alegría» e «ira»), las proporciones no diferían significativamente.

Posteriormente procedimos a analizar las tasas de reconocimiento de los estímulos en función del nivel de alexitimia, para ello se realizó un ANOVA de dos factores 3×2 (GRUPO DE ALEXITIMIA X SEXO) sobre la variable «número de errores cometidos en la tarea de identificación», tanto para el conjunto de los ocho estímulos, como para las cuatro primeras y las cuatro últimas frases-estímulo. Se decidió realizar sendos análisis, fraccionando las cuatro primeras y las cuatro últimas frases, puesto que en estudios anteriores (Montero et al., 1998) se observó un efecto atribuido a la práctica en la tasa de reconocimiento de los estímulos. Los resultados mostraron la ausencia de diferencias significativas para cada uno de los factores, tanto para la variable Sexo ($F[1] = .117$; $p = .733$), como para el Nivel de Alexitimia ($F[2] = .677$; $p = .510$), así como en su interacción ($F[2] = .894$; $p = .411$). Es preciso señalar que la variable dependiente «número de reconocimiento de las cuatro primeras frases» tiende a la significación estadística ($F[2] = 2.817$; $p = .063$), en relación al factor Grupo de Alexitimia.

Finalmente, se analizó el número de frases no identificadas por los sujetos (respuestas señaladas en el apartado «no sé») durante la tarea, tanto para el total de los estímulos como para las cuatro primeras y cuatro últimas frases. Los resultados muestran la existencia de diferencias significativas en el factor *Grupo de Alexitimia* en la variable «Número de frases no reconocidas en los cuatro primeros estímulos» ($F[2]= 4.994; p<.01$). No se observó ningún otro efecto significativo.

La existencia de diferencias significativas en el número de frases no reconocidas en los cuatro primeros estímulos puede apreciarse gráficamente en la Tabla 3 y la Figura 1, donde se muestra cómo el nivel de alexitimia de los sujetos se corresponde con un incremento en el número de estímulos no reconocido en función del nivel de alexitimia.

Frases	Nivel de alexitimia		
	Bajo	Medio	Alto
Total frases no reconocidas	.78 (1.12)	.79 (1.29)	1.24 (1.77)
No reconocidas en las 4 primeras frases	.33 (.73)	.40 (.68)	.81 (.91)
No reconocidas en las 4 últimas frases	.40 (.69)	.39 (.75)	.46 (.80)

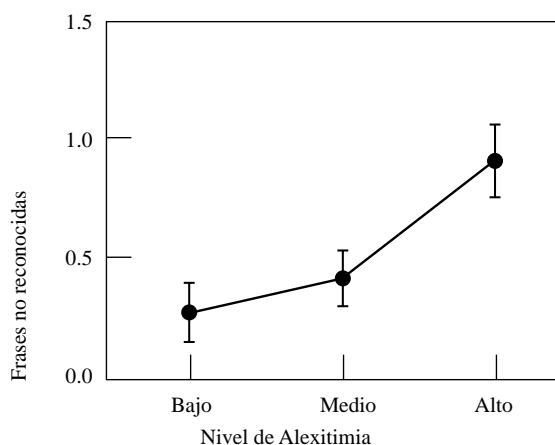


Figura 1. Número de frases no reconocidas en los cuatro primeros estímulos, agrupadas por el nivel de alexitimia de los sujetos

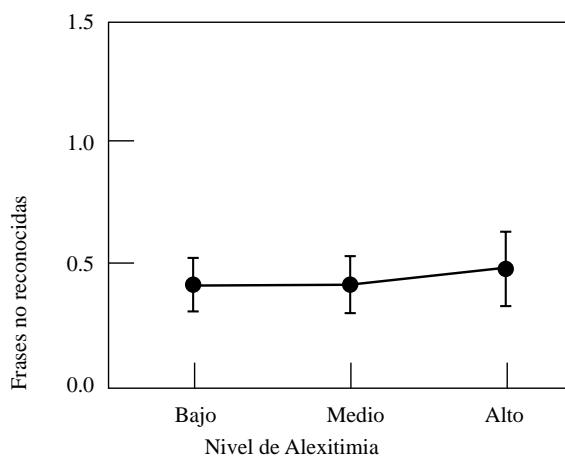


Figura 2. Número de frases no reconocidas en los cuatro últimos estímulos, agrupadas por el nivel de alexitimia de los sujetos

Por el contrario, en las cuatro últimas frases, el número de frases no reconocidas no guarda relación alguna con el nivel de alexitimia (véase la Figura 2).

Discusión y conclusiones

Los resultados obtenidos confirman parcialmente la hipótesis principal de este estudio, mostrando que los sujetos con altos niveles de alexitimia identifican significativamente menos expresiones emocionales que quienes poseen bajos niveles, durante la presentación de la primera mitad de los estímulos. Este efecto puede atribuirse al efecto de la práctica, tal y como se ha observado en estudios anteriores (Montero et al., 1998).

Las tasas de reconocimiento obtenidas con el empleo de voz sintética son sensiblemente inferiores a las informadas por otros autores (Schröder, 1999), así como a las obtenidas en el estudio inicial realizado para valorar la validez de los estímulos tras el proceso de síntesis (Montero et al., 1998), debido, posiblemente, a que en el presente trabajo empleamos ocho estímulos frente a los veinte utilizados para validar estos estímulos. La intención de dificultar la tarea de reconocimiento a los sujetos aconsejó reducir el número de estímulos para minimizar efectos debidos a la práctica a lo largo del desarrollo de la tarea.

Al igual que el estudio anteriormente citado, identificamos un efecto significativo en la tasa de reconocimiento de las cuatro últimas frases, que atribuimos a la práctica. De ellas, la expresión de tristeza es la más reconocida (en el 50% de las ocasiones aproximadamente), mientras que en el polo opuesto se sitúa la frase que expresa alegría, reconocida entre el 37 y el 42 por ciento de las ocasiones. Indudablemente esta tasa es muy reducida si tenemos en cuenta que la simple respuesta al azar arrojaría tasas de reconocimiento en torno al 25%. Al igual que en la mayoría de los estudios, la tristeza y la ira son las expresiones mejor reconocidas (Cahn, 1990), mientras que la alegría es la que más dificultades de reconocimiento entraña (Wallbott y Scherer, 1986).

Estas tasas de reconocimiento son inferiores a las habitualmente obtenidas cuando los estímulos son codificados por actores; así, de los estudios revisados por Scherer y Scherer (1981), la media de reconocimiento se sitúa en torno al 60%. En general, existen evidencias de que ciertas emociones (*alegría, sorpresa e ira*) son mejor reconocidas por la expresión facial, mientras que otras (*miedo y tristeza*) lo son por la vocal; en el caso del *asco*, parece reconocerse mejor cuando la información de que dispone el sujeto combina ambas modalidades sensoriales (de Silva, Miyasato y Nakatsu, 1997).

El hecho de que los sujetos con altos niveles de alexitimia reconozcan significativamente un menor número de expresiones supone un apoyo a la hipótesis cognitiva de la alexitimia; sin embargo, los resultados no muestran relación alguna entre el nivel de alexitimia y el número de errores de reconocimiento, si bien los resultados rondan la significación estadística. A nuestro juicio, este hecho no desvirtúa sensiblemente el valor de nuestros resultados, por cuanto cabe atribuir cualidades distintas a ambas variables dependientes. Mientras que «el número de errores cometidos» por los sujetos implica para éstos la emisión de un juicio erróneo, al que los sujetos son renuentes en tareas en que se evalúa rendimiento, por el contrario, la variable dependiente «frases no reconocidas» representa una modalidad de respuesta más conservadora, que refleja más adecuadamente, a nuestro juicio, las alteraciones presentes en la alexitimia en relación a la identificación de estímulos afectivos,

ya que conceptualmente se entiende que ésta refleja una alteración en la identificación (expresada en este estudio en la respuesta «no sé») más que la inadecuada categorización de los mismos (reflejada en el número de errores). Sin embargo, la gran variabilidad presente en las tasas de reconocimiento de los grupos induce a pensar la influencia de otras variables en estos resultados.

Estos resultados, conjuntamente con los obtenidos en otras tareas en que se han evaluado la existencia de sesgos en el procesamiento de estímulos de carácter afectivo, suponen una contribución a la hipótesis cognitiva de la alexitimia. A nuestro juicio la manera más adecuada de entender el trastorno pasa por la consideración del trastorno como una alteración multicausal, entre los que se incluyen influencias genéticas, socioculturales, alteraciones neurobiológicas, deficiencias en el desarrollo afectivo, así como déficits en el procesamiento cognitivo de la información y regulación afectiva (Martínez-Sánchez y Ortiz, 2001).

Finalmente, en el ámbito de las aplicaciones de los procedimientos de CTV, y más específicamente las telecomunicaciones, estos resultados son potencialmente relevantes para explicar, en parte, las diferencias individuales en el reconocimiento de los es-

tímulos generados por estos procedimientos, junto a otras muchas variables de índole atencional, perceptiva y afectiva. Progresivamente los estímulos vocales de naturaleza sintética a los que inexorablemente estamos sometidos se incrementa día a día, dadas sus numerosas aplicaciones, entre las que destacamos: 1) utilidades para invidentes que posibilitan leer una publicación en formato digital o el contenido de la pantalla de un monitor de ordenador con aplicaciones ofimáticas o de Internet; 2) lectura de mensajes de correo electrónico a través de la línea telefónica empleando terminales analógicos o digitales de banda estrecha; 3) agentes animados o ayudantes personales digitales (*Digital Personal Assistant*), en los que los interfaces hombre-máquina, cada vez más amigables, incorporan capacidades de inteligencia y comunicación emocional; aplicaciones inconcebibles si carecen de capacidad de CTV y reconocimiento de habla; 4) servicios vocales interactivos telefónicos, automáticos o semiautomáticos para la obtención de información, en los que la síntesis de voz permite cambiar los mensajes del servicio sin necesidad de volver a grabarlos; estos servicios permiten al usuario realizar reservas de billetes, consultas de saldos, compras por teléfono, etc.

Referencias

- Bagby, R.M., Parker, J.D.A., Taylor, G.J. y Acklin, M.W. (1993). Alexithymia and the ability to distinguish different emotional states. *Psychotherapy and Psychosomatics*, 55, 122.
- Bagby, R.M., Parker, J.D. y Taylor, G.J. (1994). The twenty-item Toronto Alexithymia Scale-I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, 38(1), 23-32.
- Bagby, R.M., Taylor, G.J. y Parker, J.D. (1994). The twenty-item Toronto Alexithymia Scale-II. Convergent, discriminant, and concurrent validity. *Journal of Psychosomatic Research*, 38(1), 33-40.
- Berenbaum, H. y Prince, J.D. (1994). Alexithymia and the interpretation of emotion-relevant information. *Emotion and Cognition*, 8(3), 231-244.
- Bradshaw, J.L. (1989). *Hemispheric specialization and psychological function*. Chichester: Wiley.
- Burkhardt, F. y Sendlmeier, F.W. (2000). Verification of acoustical correlates of emotional speech using formant-synthesis. *ISCA Workshop on Speech and Emotion*, Belfast.
- Cahn, J.E. (1990). *Generating expression in synthesized speech*. Technical Report of MIT, Media Laboratory. Massachusetts: MIT.
- De Silva, L.C., Miyasato, T. y Nakatsu, R. (1997). Facial emotion recognition using multimodal information. *Proceedings of the International Conference on Information, Communication and Signal Processing (ICICS '97)*, Singapore, (pp. 397-401).
- Dewaraja, R. y Sasaki, Y. (1990). A left to right hemisphere callosal transfer deficit of nonlinguistic information in alexithymia. *Psychotherapy and Psychosomatics*, 54(4), 201-207.
- Infrasca, R. (1997). Alexithymia, neurovegetative arousal and neuroticism. An experimental study. *Psychotherapy and Psychosomatics*, 66, 276-280.
- Jessimer, M. y Markham, R. (1997). Alexithymia: a right hemisphere dysfunction specific to recognition of certain facial expressions? *Brain and Cognition*, 34, 246-258.
- Klatt, D. (1987). Review of text-to-speech conversion for English. *Journal of the Acoustical Society of America*, 82(3).
- Lamberty, G.J. y Holt, G.S. (1995). Evidence for a verbal deficit in alexithymia. *Journal of Neuropsychiatry*, 7(3), 320-324.
- Lumley, M.A., Stettner, L. y Wehmer, F. (1996). How are alexithymia and physical illness linked? A review and critique of pathways. *Journal of Psychosomatic Research*, 41(6), 505-518.
- Lumley, M.A., Tomakowsky, J. y Torosian, T. (1997). The relationship of alexithymia to subjective and biomedical measures of disease. *Psychosomatics*, 38, 497-502.
- Mandal, M.K. y Singh, S.K. (1990). Lateral asymmetry in identification and expression of facial emotions. *Cognition and Emotion*, 4, 61-70.
- Martin, J. B. y Pihl, R.O. (1985). The stress-alexithymia hypothesis: theoretical and empirical considerations. *Psychotherapy and Psychosomatics*, 43, 169-176.
- Martínez-Sánchez, F. (1996). Adaptación española de la Escala de Alexitimia de Toronto (TAS-20). *Clínica y Salud*, 7(1), 19-32.
- Martínez-Sánchez, F. (1999). La alexitimia: un factor de riesgo para el padecimiento de los efectos patógenos del estrés. In E.G. Fernández-Abascal y F. Palmero (Eds.), *Emoción y Salud* (pp. 387-401). Barcelona: Ariel.
- Martínez-Sánchez, F. y Fernández Castro, J. (1994). Emoción y Salud. Desarrollos en Psicología Básica y Aplicada. *Anales de Psicología*, 10(2), 101-109.
- Martínez-Sánchez, F. y Marín, J. (1997). Influencia del nivel de alexitimia en el procesamiento de estímulos emocionales en una tarea Stroop. *Psicothema*, 9(3), 519-527.
- Martínez-Sánchez, F. y Ortiz, B. (2000). La evaluación de la alexitimia. In M. Casullo y Páez, D. (Ed.), *Alexitimia y cultura* (pp. 35-51). Buenos Aires: Paidós.
- Martínez-Sánchez, F. y Ortiz, B. (2001). Consideraciones en torno a la etiología de la alexitimia. *Boletín de Psicología*, 70, 11-28.
- Martínez-Sánchez, F., Ato, M., Córcoles, E., Huedo, T. y Selva, J. (1998). Stability in the alexitimia levels: a longitudinal analysis of temporary series on various emotional answers. *Personality and Individual Differences*, 24(6), 767-772.
- Martínez-Sánchez, F., Ortiz, B. y Ato, M. (2001). Subjective and autonomic stress responses in alexithymia. *Psicothema*, 13(1), 63-68.
- McDonald, P.W. y Prkachin, K.W. (1990). The expression and perception of facial emotion in alexithymia: a pilot study. *Psychosomatic Medicine*, 52, 199-210.
- Montero, J.M., Gutiérrez-Arriola, J., Colás, J. y Pardo, J.M. (1999). Analysis and Modelling of Emotional Speech in Spanish. *Proceedings of the XIVth International Congress of Phonetic Sciences*, Vol. II, pp. 957-960, San Francisco.
- Montero, J.M., Gutiérrez-Arriola, J., de Córdoba, R., Enríquez, E. y Pardo, J.M. (2001). The role of pitch and tempo in Spanish emotional speech: towards concatenative synthesis. In E. Keller y G. Bailey (Eds.), *Improvements in speech synthesis*. John Wiley & Sons, 2001.
- Montero, J.M., Gutiérrez-Arriola, J., Palazuelos, S., Enríquez, E., Aguilera, S. y Pardo, J.M. (1988). Emotional Speech Synthesis: from Speech

- Database to TTS. *Proceedings of the International Conference in Speech and Language Processing ICSLP'98*, Vol. III, 923-926. Sydney, Australia.
- Montreuil, M. y Jouvent, R. (1989). Le test de la saisie visuelle parallèle. *Encéphale*, 15, 409-413.
- Montreuil, M., Jouvent, R., Carton, S., Bungener, C. y Widlocher, D. (1991). Parallel Visual Information Processing Test. An experimental assessment of alexithymia. *Psychotherapy and Psychosomatics*, 56(4), 212-219.
- Murray, I.R. y Arnott, J.L. (1995). Implementation and testing of a system for producing emotion-by-rule in synthetic speech. *Speech Communication Journal*, 16, 359-368.
- Murray, I.R. y Arnott, J.L. (1996). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *Journal of the Acoustical Society of America*, 93(2), 1.097-1.108.
- Näring, G.W. y van der Staak, C.P. (1995). Perception of heart rate and blood pressure: the role of alexithymia and anxiety. *Psychotherapy and Psychosomatics*, 63, 193-200.
- Páez, D., Martínez-Sánchez, F., Velasco, C., Mayordomo, S., Fernández, I. y Blanco, A. (1999). Validez psicométrica de la Escala de Alexitimia de Toronto: un estudio transcultural. *Boletín de Psicología*, 63, 55-76.
- Parker, J.D., Bagby, R.M., Taylor, G.J., Endler, N.S. y Schmitz, P. (1993). Factorial validity of the 20-item Toronto Alexithymia Scale. *European Journal of Personality*, 7, 221-232.
- Parker, J.D., Taylor, G.J. y Bagby, R.M. (1993). Alexithymia and processing of emotional stimuli: an experimental study. *New Trends in Experimental and Clinical Psychiatry*, IX(1/2), 9-14.
- Rank, E. y Pirker, H. (1998). Generating Emotional Speech with a Concatenative Synthesiser. (1998). *Proceedings of the International Conference on Spoken Language Processing*, 3, 671-674.
- Roedema, T.M. y Simons, R.F. (1999). Emotion-processing deficit in alexithymia. *Psychophysiology*, 36, 379-387.
- Sachse, R. (1994). Herzsclangwahrnehmung bei psychosomatischen Patienten: abwendung der aufmerksamkeit von eigenen körperprozessen. *Psychotherapie, Psychosomatik Medizinische Psychologie*, 44(8), 284-292.
- Scherer, K.R. y Scherer, U. (1981). Speech behavior and personality. En J. Darby (Ed.), *Speech evaluation in psychiatry* (pp. 115-135). New York: Grune & Stratton.
- Schröder, M. (1999). Can emotions be synthesized without controlling voice quality. *Phonus*, 4, 37-55.
- Sifneos, P.E. (1973). Prevalence of «alexithymic» characteristics in psychosomatic patients. *Psychotherapy and Psychosomatics*, 22, 255-262.
- Silberman, E.K. y Weingartner, H. (1986). Hemispheric lateralization of functions related to emotion. *Brain and Cognition*, 5, 322-353.
- Suslow, T. (1998). Alexithymia and automatic affective processing. *European Journal of Personality*, 12, 433-443.
- Taylor, G.J. (2000). Recent developments in alexithymia theory and research. *Canadian Journal of Psychiatry*, 45(2), 134-142.
- Taylor, G.J., Bagby, R.M. y Parker, J.D. (1991). The alexithymia construct. A potential paradigm for psychosomatic medicine. *Psychosomatics*, 32(2), 153-164.
- Taylor, G.J., Bagby, R.M. y Parker, J.D. (1997). *Disorders of affect regulation. Alexithymia in medical and psychiatric illness*. Cambridge: Cambridge University Press.
- Wallbott, H.G. y Scherer, K.R. (1986). Cues and channels in emotion recognition. *Journal of Personality and Social Psychology*, 51, 690-699.