

# Círculo de Lingüística Aplicada a la Comunicación

ISSN: 1576-4737

 EDICIONES  
COMPLUTENSE<http://dx.doi.org/10.5209/clac.70562>

## Las emociones básicas en un cuento: percepción audiovisual de los niños

Aintzane Etxebarria Lejarreta<sup>1</sup>; Iñaki Gaminde Terraza<sup>2</sup>; Asier Romero Andonegi<sup>3</sup>; Aitor Iglesias Chaves<sup>4</sup>

Recibido: 10 de septiembre de 2018 / Aceptado: 28 de mayo de 2020

**Resumen.** El objetivo de este trabajo es determinar si la elección lingüística coincide con la visual al escuchar una emoción. Para ello se grabó el cuento de Caperucita Roja en el que aparecían tres emociones básicas simuladas por una joven bilingüe procedente del País Vasco y se diseñó un test específico para la investigación, en el que los 178 participantes bilingües con edades comprendidas entre los 8 y 15 años, (99 son de género masculino y el 79 son del femenino), estudiantes de tercer y cuarto curso de Educación Primaria y de primero y segundo de Educación Secundaria Obligatoria, de un colegio concertado tenían que elegir la emoción básica a la que correspondía lo escuchado. Según los resultados, vincular el estímulo auditivo con el tipo de emoción verbalizada no es complicado, genera contradicciones y parece que entraña una mayor dificultad para los sujetos relacionar éstos con los estímulos visuales.

**Palabras clave:** emociones básicas, percepción audiovisual, prosodia.

### [en] The basic emotions in a tale: audiovisual perception in children

**Abstract.** The objective of this paper is to determine if linguistic selection coincides with visual selection when hearing an emotion. For that purpose, the story of Little Red Riding Hood was recorded, in which three basic emotions appeared, simulated by a young bilingual girl from the Basque Country. A specific test was designed for the study, in which 178 bilingual participants between the ages of 8 and 15 (99 male and 79 female), students of the third and fourth course of Primary Education and of the first and second in Compulsory Secondary Education from a state-assisted private school, had to choose the basic emotion that corresponded to what was heard. According to the results, linking the auditory stimulus relationship with the type of emotion verbalised is not complicated, creates contradictions, and seems to entail more difficulty for the subjects to link these with the visual stimuli

**Keywords:** basic emotions, audiovisual perception, prosody

**Cómo citar:** Etxebarria, A., Gaminde, I., Romero, A. y Iglesias, A. (2020). Las emociones básicas en un cuento: percepción audiovisual de los niños. *Círculo de Lingüística Aplicada a la Comunicación* 83, 41-50, <http://dx.doi.org/10.5209/clac.70562>

**Índice.** 1. Introducción teórica. 2. Metodología. 2.1. Construcción del test. 2.1.1. Análisis de contenido. 2.1.2. Redacción de los ítems. 2.2. Participantes. 2.3. Procedimiento. 2.4. Análisis de los datos. 3. Conclusiones y prospectiva de futuro. Agradecimientos. Bibliografía

## 1. Introducción teórica

La prosodia es un conjunto de rasgos suprasegmentales en los que se incluyen el acento, la entonación, el tono, el ritmo, la melodía, las pausas, la velocidad de elocución y la cualidad de la voz (Llisterri, 2016), y es un rasgo de la lengua oral que ayuda a llevar estados emocionales o afectivos en las oraciones (Saá, 2001). Numerosos autores distinguen entre prosodia semántica o lingüística y afectiva o emocional (Couper-Kuhlen, 1986; Murphy y Cutting, 1990; Edwards et al., 2001; Van Rijn et al., 2005; Hoekert et al., 2007; Difaldis et al., 2013); Martínez et al. (2015) definen ambos conceptos de la siguiente manera:

Estos dos tipos de prosodia son definidos como: prosodia semántica, que es la que se refiere al significado, y prosodia afectiva, que es la que se refiere a la expresión de las emociones. Son estos tres parámetros, junto con la gestualidad y el contenido del discurso, los que nos dan la sensación de emoción cuando escuchamos a alguien, y de acuerdo con estudios anteriores, podemos asociar ciertos parámetros acústicos con el reconocimiento y la expresión de emociones. (Martínez et al., 2015: 14).

<sup>1</sup> UPV/EHU. Correo electrónico: aintzane.etxebarria@ehu.eus

<sup>2</sup> UPV/EHU. Correo electrónico: inaki.gaminde@gmail.com

<sup>3</sup> UPV/EHU. Correo electrónico: a.romero@ehu.eus

<sup>4</sup> UPV/EHU. Correo electrónico: aitor.iglesias@ehu.eus

El hablante debe desarrollar una adecuada competencia prosódica que se define como la capacidad para emplear la voz de manera expresiva, coherente y adecuada al contexto (Mozziconacci, 1998; Grau i Tarruell y Vilá i Santasusana, 2009; Etxebarria, 2010; Saá, 2011; Garay y Gaminde et al., 2011, 2012a, 2012b) sin dejar de lado la capacidad para percibir, sobre todo, aquellos mensajes que tienen carga emotiva o afectiva. En la literatura científica del campo de la medicina, y más concretamente en el área de la psiquiatría, se han realizado numerosos trabajos sobre los rasgos de la prosodia afectiva o emocional con los pacientes con enfermedades mentales como la esquizofrenia (Murphy y Cutting, 1990; Leentjens et al., 1998; Ross et al., 2001; Matsumoto, 2006; Pijnenborg, G.H., 2007), estas investigaciones revelan que estos enfermos, al igual que aquellos que tienen el hemisferio derecho dañado, tienen un deterioro de la percepción y expresión de la prosodia emocional (Hoekert et al., 2007) denominado aprosodia, puesto que el habla es monótono y no muestra ningún tipo de estado afectivo ni emotivo (Leon y Rodríguez, 2008; Saá, 2011). Por tanto, el habla emocional lo podemos definir de la siguiente manera:

An episode of interrelated, synchronized changes in the states of all or most of the five organismic subsystems in response to the evaluation of an external or internal stimulus event as relevant to major concerns of the organism. (Scherer, 2005).

En la literatura científica sobre el estudio de las emociones (Liscombe, 2007; Gaminde, 2010), se ha subrayado la dificultad metodológica para recoger el habla emocional en contextos naturales, razón por la que se analizan las emociones actuadas. Sin embargo, han sido numerosas las investigaciones que se han publicado en los últimos años sobre la expresión y percepción de las distintas emociones, muestra de ello es la variedad de autores que citan Brøndsted, Nielsen, y Ortega (2000).

Most studies concentrate on between four and seven different emotions, e.g. neutrality, happiness, anger, sadness in (Carlson et al. 1992) and in (Montero et al. 1998), neutrality, surprise, happiness, anger, sadness in (Engberg et al. 1997), neutrality, happiness, hot anger, cold anger, sadness in (Pereira et al. 1998), neutrality, joy, boredom, anger, sadness, fear, indignation in (Mozziconacci 1998). (Brøndsted, Nielsen, y Ortega, 2000).

Cada emoción se define acústicamente por unos parámetros concretos (Martínez et al., 2012); Mozziconacci (1998, 2001) define los valores óptimos de los parámetros prosódicos para la transmisión de seis emociones en holandés que son la alegría, el aburrimiento, el enojo, la tristeza, el miedo y la indignación. Otros trabajos como el de Gustafson-Capková (2001) subrayan la importancia de las emociones en la creación de las tecnologías del habla. En el ámbito de la percepción varios estudios dan a entender la importancia de la entonación en la identificación de las emociones (Soskin y Kauffman, 1961; Brown, 1980; Bance y Scherer, 1996; Mozziconacci, 1998; Pereira, 2000; Scherer, 2005). Finalmente, Liscombe (2007) analiza características prosódicas de las emociones y su percepción, señalando la importancia de los componentes suprasegmentales del habla (tono, ritmo y volumen) como características importantes que definen al hablante.

En la lengua vasca se han definido y analizado la producción y la percepción de las emociones con diferentes objetivos; el de crear el habla sintética (Navas et al., 2004, 2005, 2007; Saratxaga et al., 2006; Sainz et al., 2008); el de desarrollar material con fines didácticos (Gaminde et al., 2014; Gaminde et al., 2015); y, por último, el de crear mapas descriptivos basados en la variación diatópica (Gaminde, 2007, 2010; Garay et al. 2011). Por lo tanto, toda esta trayectoria ha tenido como resultado la definición de seis emociones en la lengua vasca (tabla 1) según las características de los parámetros acústicos (largo/corto, pequeño/grande):

**Tabla 1.** Características de los parámetros acústicos en cada emoción. (Gaminde et al., 2014).

Emoción	Duración	Rango de F0	Rango de Energía
Miedo	Larga	Pequeño	Pequeño
Asco	Larga	Pequeño	Pequeño
Tristeza	Larga	Pequeño	Pequeño
Enfado	Corta	Grande	Grande
Asombro	Corta	Grande	Grande
Alegría	Corta	Grande	Grande

En cuanto a las características gestuales de las emociones, se han analizado en diferentes ámbitos de estudio, como en el campo de la psiquiatría (Freedman y Hoffman, 1967; McQuown et al., 1971), o en el de la lingüística y la antropología (Pike, 1967; Ekman y Friesen, 1969; Birdwhistell, 1970; Efron, 1972; Kendon, 1972; MacNeill, 1992; Loehr, 2004; Esteve-Gibert y Prieto, 2013). Ekman et al. (2002) tienen en cuenta la posición de la cabeza, los ojos, los labios, la mandíbula y las cejas, y definen un sistema de codificación llamado *Facial Action Coding System*, a pesar de que en un principio se pensaba que no eran factores universales sino cambiantes que se aprendían según las normas sociales:

In studies in New Guinea, Borneo, The United States, Brazil, and Japan we found evidence of pan-cultural elements in facial displays of affect. Observers in these cultures recognize some of the same emotions when they are shown a standard set of facial photographs. This finding contradicts (i) the theory (1) that facial displays of emotion are socially learned and therefore culturally variable; and (ii) the finding from studies within a single cultura that observers of the face alone do not achieve either accuracy or high agreement in recognizing different emotional states. (Ekman et al., 1969:86).

A día de hoy, no está muy clara la relación entre los gestos y la entonación para la decodificación del mensaje, pero Crespo Sendra et al. (2013) hacen referencia a la investigación de Srinivasan y Massaro (2003) y mencionan la teoría *Fuzzy Logical Model of Perception* (Massaro, 1989) para señalar que el componente sonoro y el visual son complementarios, y uno de ellos ayuda cuando el otro es ambiguo:

For example, Srinivasan and Massaro (2003) performed five perception experiments to investigate the role of facial expression and intonation in distinguishing statements from echo questions in English. They concluded that statements and questions are distinguished both intonationally and visually but that intonational cues have greater influence as compared to the visual cues. Their data were consistent with the Fuzzy Logical Model of Perception (FLMP, Massaro, 1989), which predicts that both auditory and visual modalities will influence the perception of prosody, and that the influence of one modality will be greater when the information provided by the other modality is ambiguous. (Crespo Sendra et al., 2013: 2)

Sidera et al. (2012) afirman fundamentados en la literatura revisada “que los niños pueden utilizar el paralenguaje vocal y facial para detectar estados emocionales en otras personas” (2012: 47). Esta circunstancia se ha evidenciado también en otros estudios que analizan esta relación teniendo en cuenta las estrategias prosódicas y la producción multimodal del niño a la hora de percibir las emociones (Esteve-Gibert & Prieto, 2014). Por tanto, en este trabajo, y partiendo de estas dos ideas, se pretende responder a la siguiente pregunta: ¿Coincide la elección lingüística con la visual al escuchar un estímulo sonoro?

## 2. Metodología

### 2.1. Construcción del test

En este apartado se especifica tanto el análisis del contenido del test como el proceso para la redacción de los ítems.

#### 2.1.1. Análisis de contenido

La primera fase consistió en la realización de un análisis de contenido inductivo (Hsieh y Shannon, 2005) sobre el Decreto 236/2015, de 22 de diciembre, por el que se establece el currículo de Educación Básica y se implanta en la Comunidad Autónoma del País Vasco. Un documento en el que se regularizan los elementos que determinan los procesos de enseñanza y aprendizaje para cada una de las enseñanzas y etapas educativas; además, se establecen los objetivos de cada enseñanza y etapa educativa atendiendo a las competencias, y se describen el conjunto de conocimientos, habilidades, destrezas y actitudes que contribuyen al logro de los objetivos tanto de cada enseñanza y etapa educativa como de las competencias trabajadas. La finalidad de este análisis es detectar las categorías relacionadas con las características que distinguen a la producción y la percepción de las emociones desde una conceptualización prosódica.

Esta relación tentativa de categorías se ha sometido a un proceso de revisión y depuración a través de un panel de profesores del centro en el que se ha realizado la prueba. Después de consultar la bibliografía pertinente sobre el tema y de realizar varias iteraciones se llegó a un acuerdo consensuado sobre el diseño y las características que integran el constructo percepción audiovisual, atendiendo a tres parámetros: (1) percepción auditiva de la emoción, (2) verbalización de la emoción, (3) elección visual de la emoción.

#### 2.1.2. Redacción de los ítems

Una vez que se ha definido el modelo sustantivo que ha de soportar el test de percepción y las dos categorías de valores que integran el constructo señaladas en el apartado anterior, se redactan los ítems relacionados con esas dos

categorías. En total, se construyeron 14 ítems que se corresponden con los 8 estímulos audiovisuales que constituyen el corpus. La duración aproximada de aplicación fue de 10 minutos. Para la producción de los estímulos se seleccionó a una alumna de la Facultad de Educación de Bilbao (UPV/EHU), matriculada en primer curso del Grado de Educación Infantil y natural de la localidad de Bermeo (Bizkaia). La selección de esta alumna estuvo motivada por la lengua materna de producción, ya que se pensó que lo más adecuado era que las personas encargadas de reproducir el cuento lo hiciesen en la variedad dialéctica del euskera hablado en Bermeo. En consecuencia, esta alumna tuvo que reproducir diferentes pasajes del cuento de Caperucita Roja (versión de los hermanos Grimm), y cada una con un rol diferente: el de narrador y el de Caperucita Roja/Lobo/Abuela. Además, la alumna tuvo que condicionar su discurso atendiendo a las distintas emociones que aparecen en la narración del cuento (alegría, miedo y sorpresa). Una vez que se preparó el guión y se realizaron diferentes pruebas se realizó la grabación final mediante el software *Audacity* (<https://www.audacityteam.org>) con micrófono externo.

Con estos estímulos auditivos se diseñó el cuestionario que debía de cumplimentar el alumnado. Este cuestionario contaba con dos partes, una primera centrada en datos de carácter general (edad, género y lengua materna); y una segunda parte con ocho ítems. En uno de ellos solamente se preguntaba por el estímulo sonoro, puesto que el protagonista era el lobo, y en los siete restantes había dos preguntas: (1) una pregunta que aludía al tipo de emoción que transmitía la grabación, es decir, los participantes tenían que marcar la emoción que percibían al escuchar la grabación (2) una pregunta que se centraba en el tipo de emoción que se reflejaba en tres rostros de mujer, y que simbolizaban el rol de Caperucita Roja o su abuela. Los rostros que expresaban sorpresa (Código: 1E+2E+5E+25C+26E), alegría (Código: 7C+12D+25D+26B) y miedo (Código: 1B+2B+4C+5B+20A+25C+26C+38A) fueron seleccionados del *Facial Action Coding System* (2002) de Ekman por su reconocimiento internacional.

## 2.2. Participantes

En este estudio han participado 178 informantes bilingües con edades comprendidas entre los 8 y 15 años, (99 son de género masculino y el 79 son del femenino), estudiantes de tercer y cuarto curso de Educación Primaria y de primero y segundo de Educación Secundaria Obligatoria, en un colegio concertado de la localidad vizcaína de Bermeo, puesto que el grupo de investigación tiene un convenio de colaboración con dicho centro. Todos los informantes están matriculados en el modelo D, enseñanza en euskera como lengua vehicular, salvo la asignatura de lengua castellana. Además, algunos de ellos hablan las dos lenguas oficiales, es decir el castellano y la lengua vasca en el ámbito familiar (67%) y el resto una de las dos lenguas (33%)

## 2.3. Procedimiento

Una vez diseñado el test, se procedió a su implementación en el centro escolar seleccionado. Previamente, hubo una reunión con el centro y con el profesorado implicado para que tuvieran conocimiento informado de la prueba. Además, se cumplimentaron los pertinentes permisos de confidencialidad para realizar dicho experimento. La prueba en sí la realizaron dos miembros del grupo de investigación y su procedimiento de aplicación fue sencillo. Antes del comienzo del experimento, se explicó al alumnado los pormenores y características del mismo. A continuación, se fueron reproduciendo cada uno de los estímulos que constituyen el corpus y la duración aproximada de aplicación fue de 10 minutos. Cada estímulo fue representado dos veces, dejando una pequeña pausa de 10 segundos entre cada uno de ellos. Después de la segunda representación, el alumnado debía seleccionar de entre las tres opciones posibles la emoción percibida en cada una de las grabaciones, y a continuación identificar dicha emoción con respecto a las tres expresiones faciales de sorpresa, alegría y miedo, y que simbolizaban el rol de Caperucita Roja o su abuela.

## 2.4. Análisis de los datos

Teniendo en cuenta las características y el tipo de test que se ha realizado presentamos el análisis de los datos dividido en tres secciones; en la primera sección se presenta el análisis acústico de los estímulos utilizados, seguidamente se analizan las respuestas dadas por los informantes a los estímulos sonoros, y finalmente se analizan las respuestas de los estímulos visuales.

Para realizar el análisis acústico de los datos se han medido la frecuencia fundamental o  $f_0$ , la energía y la duración total de los estímulos. En lo que respecta al  $f_0$  se consideran dos medidas diferentes, la media total de  $f_0$  del estímulo ( $X_{f_0}$ ) y el rango de  $f_0$  ( $H_{f_0}$ ) ambas en hercios (hz); este mismo criterio se ha utilizado en lo que respecta a la energía de los estímulos, es decir, se han considerado la media de la energía total de los estímulos ( $X_{db}$ ) y el rango de la misma ( $H_{db}$ ), ambas medidas se dan en decibelios (db). En cuanto a la duración sólo se ha considerado la duración total de cada estímulo en milisegundos (ms).

En la tabla 2 se presentan los resultados de las mediciones para cada característica acústica considerada en los estímulos.

**Tabla 2.** Características acústicas de cada emoción.

Estímulo	Xf0	Hf0	Xdb	Hdb	ms
1	263,02	225,84	57,86	38,44	1210
2	255,96	261,45	55,94	39,81	1730
3	347,61	189,21	62,14	40,80	1540
4	332,22	157,22	65,11	38,76	1110
5	223,77	167,33	54,87	48,13	1160
6	281,29	199,36	56,19	41,57	1180
7	284,40	174,87	63,83	37,64	720
8	299,75	296,57	64,39	41,99	1240

Una vez aplicado el test se procedió definir cada estímulo desde la percepción del alumnado (Tabla 3).

**Tabla 3.** Tipo de estímulo desde la percepción del alumnado.

Estímulo	Alegría		Miedo		Sorpresa	
	Cantidad	%	Cantidad	%	Cantidad	%
1	175	98,3%	1	0,56%	1	0,56%
2	101	56,74%	2	1,12%	75	42,13%
3	0	0%	166	93,25%	12	6,74%
4	176	98,8%	0	0%	2	1,12%
5	4	2,24%	9	5,05%	165	92,69%
6	1	0,5%	22	12,35%	154	86,51%
7	3	1,68%	136	76,40%	38	21,34%
8	163	91,5%	3	1,68%	12	6,74%

Si analizamos las características acústicas de los estímulos elegidos por los informantes podemos elaborar la tabla 4 con las medias y desviaciones de las mismas.

**Tabla 4.** Características acústicas de los estímulos elegidos por los informantes.

	Alegría		Miedo		Sorpresa	
	$\bar{x}$	des	$\bar{x}$	des	$\bar{x}$	des
Xf0	298,33	34,62	316	44,69	252,53	40,67
Hf0	226,54	69,68	182,04	10,13	183,35	22,65
Xdb	62,45	3,99	62,98	1,19	55,53	0,93
Hdb	39,73	1,96	39,22	2,24	44,85	4,64
Duración	1186,67	68,07	1130	579,83	1170	14,14

Con respecto a la media del f0 observamos que el miedo es la emoción con una media más alta seguida de la alegría y la sorpresa. El análisis ANOVA da como resultado que estas diferencias son estadísticamente significativas ( $F = (gl: 2) 5,699, p = 0,006$ ), aunque realizando la prueba post hoc de Tukey comprobamos que esta diferencia sólo es válida para distinguir entre el miedo y la sorpresa ( $p = 0,004$ ). La mayor amplitud del rango de f0 se realiza con la alegría, mientras que con el miedo y la sorpresa la diferencia es muy pequeña; los resultados de ANOVA muestran que estas diferencias son estadísticamente significativas ( $F = (gl: 2) 4,833, p = 0,012$ ), la

prueba de Tukey nos da una diferencia significativa entre la alegría y el miedo ( $p = 0,017$ ). En el análisis de la energía total de los estímulos se constata que la alegría y el miedo tienen las energías más altas siendo la sorpresa la que presenta la media más baja. El ANOVA muestra que esta diferencia es estadísticamente significativa ( $F = (gl: 2) 13,124, p = 0,000$ ), aunque, según la prueba post hoc de Tukey sólo es válida para distinguir entre la sorpresa y la alegría ( $p = 0,000$ ) y ésta y el miedo ( $p = 0,000$ ). La mayor amplitud del rango de energía se realiza con la sorpresa, mientras que entre la alegría y el miedo las diferencias son muy pequeñas en comparación. La prueba de ANOVA da como resultado que esta diferencia es estadísticamente significativa ( $F = (gl: 2) 9,370, p = 0,000$ ), aunque según la prueba post hoc de Tukey ésta sólo es válida para diferenciar la sorpresa de las otras dos emociones ( $p = 0,002$ ). Es interesante constatar que a pesar de que la sorpresa presenta el nivel más bajo en lo que respecta a la media de la energía a su vez tiene un mayor rango de energía que las otras dos emociones. Con respecto a la duración total de cada estímulo constatamos que las diferencias que se observan en la tabla 3 no son estadísticamente significativas.

Hemos querido analizar las respuestas de los sujetos de acuerdo con la edad, género y lengua materna de los mismos. Al realizar este análisis, hemos hallado diferencias significativas en cuanto a la variable edad solamente en el ítem 7 ( $p = 0,01$ ) que corresponde con la elección visual del estímulo sonoro. Con respecto a la variable independiente género no se dan diferencias significativas. Por último, en relación a la lengua materna (vasco, castellano o ámbas) de los jóvenes encuestados se dan diferencias estadísticamente significativas en los ítems 8 ( $p = 0,001$ ) y 14 ( $0,001$ ) que coinciden con la verbalización de la emoción.

Los jóvenes que respondieron al test, además de responder a los ítems de percepción tuvieron que identificar la emoción elegida en uno de los rostros seleccionados de *Facial Action Coding System* (2002) de Ekman, los resultados fueron los que se muestran en la tabla 5:

**Tabla 5.** Respuestas a los ítems visuales.

Ítem	Emoción elegida	Porcentaje	Emoción definida según Ekman (2002)
2	Contento	1,1	Contento
	Asustado	97,8	
5	Contento	56,7	Asustado
	Asustado	1,1	
	Asombrado	41,6	
7	Contento	1,7	Contento
	Asustado	94,9	
	Asombrado	1,1	
9	Contento	42,1	Asombrado
	Asustado	1,7	
	Asombrado	54,5	
11	Contento	46,6	Asustado
	Asustado	0,6	
	Asombrado	51,7	
13	Contento	54,5	Asombrado
	Asustado	2,2	
	Asombrado	41,6	

Como se puede observar en la tabla 5 no coincide la elección visual del alumnado con la emoción definida por Ekman (2002), de la misma manera que la percepción auditiva en la mayoría de los casos no coincide con la elección visual (tabla 6), a excepción de los ítems 5 y 6.

**Tabla 6.** Comparación de las respuestas a los ítems de percepción y señalización visual.

Estímulo	Cantidad	Porcentaje	Percepción auditiva	Señalización visual	Porcentaje
1	175	98,3%	Alegría	Asustado	97,8%
2	101	56,74%	Alegría	–	
3	166	93,25%	Miedo	Contento	56,7%
4	176	98,8%	Alegría	Asustado	94,9%
5	4	92,69%	Sorpresa	Asombrado	54,5%
6	1	86,51%	Sorpresa	Asombrado	52,7%
7	3	76,40%	Miedo	Contento	54,5%
8	163	91,5%	Alegría	–	

### 3. Conclusiones y prospectiva de futuro

El análisis de los datos sobre la identificación auditiva y clasificación de las emociones presentado en el apartado anterior muestra un alto grado de coincidencia en las respuestas de los sujetos y gran cohesión entre los mismos, ya que no hay diferencias estadísticamente significativas en lo que respecta a las variables edad, género y lengua materna de los mismos. En cambio, hemos visto que la relación entre los estímulos auditivos y visuales que se realiza por la mediación del léxico presenta un gran número de contradicciones, es decir, mientras que la relación estímulo auditivo con el tipo de emoción verbalizada resulta ser una tarea sencilla, relacionar éstos con los estímulos visuales genera contradicciones y parece que entraña una mayor dificultad para los sujetos. Asimismo, debe tenerse en cuenta que se les ha solicitado realizar tres tipos diferentes de tareas:

- (1) Identificar los estímulos auditivos
- (2) Relacionar por medio del léxico cada estímulo auditivo con una emoción
- (3) Relacionar los estímulos auditivos con los estímulos visuales por medio del léxico

Creemos que para futuros trabajos sería interesante omitir la mediación léxica y realizar el experimento tratando de relacionar los estímulos auditivos directamente con los visuales. En este trabajo nos encontramos con una serie de problemas teóricos y metodológicos añadidos que creemos pueden ser de interés en posteriores investigaciones. Por una parte, tenemos el problema de determinar un sistema formal de representación de los datos acústicos en relación con las emociones que posibilite la comparación entre ellas con datos obtenidos de hablantes diferentes. Teniendo en cuenta la variación fonética intra-sujetos e inter-sujetos, ya que, la correlación entre las medidas no es exacta, sino que está dentro de un rango en todos los casos, nos encontramos ante el problema de las diferencias de rango intra-sujetos y de la variación y no exactitud de las mismas inter-sujetos. Por último, las medidas que se obtienen de las diferentes características que forman parte del subsistema fonético de la prosodia se hallan mediatizadas por las interacciones con otros sistemas de destrezas y habilidades como el sistema léxico o el que relaciona los estímulos visuales con los auditivos.

Proponemos que la relación entre las características acústicas estudiadas y las emociones, al no ser ésta lineal, se puede representar formalmente por medio de una matriz donde las filas estarían constituidas por cada emoción y en las columnas aparecerían las características acústicas estadísticamente significativas (en este caso todas excepto la duración). Dado que la relación entre emociones y características acústicas no es exacta y depende del rango de los hablantes, éstas pueden representarse de acuerdo con el ranquin de emisión (Gaminde, 2010), en nuestro caso este ranquin se representa de 1 a 3 de mayor a menor. En la tabla 7 se muestran los vectores de cada emoción de acuerdo con el ranquin del emisor de nuestros estímulos.

**Tabla 7.** Ranquin de las emociones según las características acústicas.

Emoción	Xf0	Hf0	Xdb	Hdb
Alegría	2	1	2	2
Miedo	1	2	1	3
Sorpresa	3	3	3	1

Este tipo de representación formal también se ha propuesto en trabajos anteriores para explicar la relación entre los subsistemas fonológico y fonético de la prosodia para dar cuenta de la entonación (Gaminde y Gandarias, 2015)

Hemos mencionado anteriormente que la relación entre características acústicas y emociones no puede ser exacta y por tanto se concreta dentro de un rango. Debemos por consiguiente considerar así mismo que esta relación es estocástica (probabilística) y desde el momento que admitimos encontrarnos ante un problema de probabilidades debemos también considerar la necesidad de aceptar unos márgenes de error.

Del mismo modo, creemos que los hablantes, tanto para la emisión como para la percepción de las emociones, almacenan unos rangos con ciertos niveles de solapamiento condicionados por otros subsistemas (pragmático, sociocultural...) que interaccionan con el subsistema fonético de la prosodia. Por tanto, estamos ante un sistema que cumple ampliamente las características de los sistemas complejos dinámicos adaptativos (Beckner et al., 2009).

## Agradecimientos

Este trabajo ha sido financiado en el marco de dos proyectos de investigación: Gobierno Vasco [IT 1028/16] y Universidad del País Vasco [GIU 22/16].

## Bibliografía

- Bance, R. & Scherer, K. (1996). Acoustic Profiles in Vocal Emotion expression. *Journal of Personality and Social Psychology* 70 (3): 614-636.
- Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., Holland, J., Ke, J., Larsen-Freeman, D., & Schoenemann, T. (2009). Language Is a Complex Adaptive System: Position Paper. *Language Learning* 59-1: 1-26.
- Birdwhistell, R. L. (1970). *Kinesics and Context: Essays on Body Motion Communication*. Philadelphia: University of Pennsylvania Press.
- Brøndsted, T., Nielsen, T.D. & Ortega, S. (2000). Classification of Emotional Attitudes in Pet-directed Speech, In DALF Proceedings. Copenhagen: CST Working Papers.
- Brown, B. L. (1980). The detection of emotion in vocal qualities, In Giles, H., Robinson, W. & Smith, P. (Ed.), *Language: Social Psychological Perspectives. Selected Papers from the First International Conference on Social Psychology and Language, held at the University of Bristol*, 237-246. Oxford: Pergamon.
- Couper-Kuhlen, E. (1986). *An Introduction to English Prosody*. London: Edward Arnold.
- Crespo Sendra, V., Kaland, C., Swerts, M. & Prieto, P. (2013). Perceiving incredulity: The role of intonation and facial gestures. *Journal of Pragmatics* 47: 1-13.
- Difalcis, M., Sampedro, M. B. & Ferreres, A. (2013). Alteraciones de la prosodia emocional y lingüística por lesiones del hemisferio derecho, in *V Congreso Internacional de Investigación y Práctica Profesional en Psicología XX Jornadas de Investigación Noveno Encuentro de Investigadores en Psicología del MERCOSUR*. Buenos Aires: Facultad de Psicología - Universidad de Buenos Aires.
- Edwards, J., Pattison, P. E., Jackson, H. J. & Wales, R. J. (2001). Facial affect and affective prosody recognition in first-episode schizophrenia. *Schizophrenia Research* 48: 235-253.
- Efron, D. (1972). *Gesture and culture*. The Hague: Mouton.
- Ekman, P., Sorenson E. R. & Friesen W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science* 164(3875): 86-88.
- Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial Action Coding System. Manual and Investigator's Guide CD-ROM*. Salt Lake City: Research Nexus.
- Ekman, P. & Friesen, W. V. (1969). The Repertoire of Nonverbal Behavior. *Semiotica*, 1(1): 49-98.
- Esteve-Gibert, N. & Prieto, P. (2013). Prosodic structure shapes the temporal realization of intonation and manual gesture movements. *Journal of Speech, Language, and Hearing Research* 56(3): 850-864.
- Esteve-Gibert, N. & Prieto, P. (2014). Infants temporally coordinate gesture-speech combinations before they produce their first words. *Speech Communication* 57: 301-316.
- Freedman, N. & Hoffman, S. P. (1967). Kinetic behavior in altered clinical states: approach to objective analysis of motor behavior during clinical interviews. *Percept Mot Skills* 24(2): 527-39.
- Gaminde, I. (2007). Intonazioaren ezaugarriak emozioen pertzepzioan lau hizkuntzatan. *Uztaro* 61: 63-84.
- Gaminde, I. (2010). *Bizkaiko Gazteen Prosodiaz: Euskaraz eta Gaztelaniaz*. Bilbo: Mendebalde Kultura Alkartea & Bizkaiko Foru Aldundia.
- Gaminde, I. & Gandarias, L. (2015). Irizpide batzuk bariazioa aztertzeo intonazioan, in Aurrekoetxea, G., Romero, A. & Etxebarria, A. (Ed.), *Linguistic Variation in the Basque an Education-I / Euskararen Bariazioa eta Bariazioaren Irakaskuntza-I*, 183-205. Bilbao: UPV/EHU.
- Gaminde, I., Etxebarria, A., Garay, U. & Romero, A. (2011). Bokatiboa eta gaitasun prosodikoa: lehen urratsak. *Uztaro* 79: 5-23.
- Gaminde, I., Etxebarria, A., Garay, U. & Romero, A. (2012a). Komunikazio gaitasunetik kultura arteko komunikazio gaitasunera: Baliabide didaktikoak, in *Hizkuntzaz Jabetzen* 209-251. Bilbao: Mendebalde Kultura Alkartea.
- Gaminde, I., Etxebarria, A., Garay, U. & Romero, A. (2012b). Esaldi barruko eten moten pertzepzioaz. *Ikastorratza* 8.



- Gaminde, I., Aurrekoetxea, G., Etxebarria, A., Garay, U. & Romero, A. (2015). Ahoskera aztergai. Ahoskera Lantzeko Argibideak eta Jarduerak. Laguntzarako materiala: Teoria eta Praktika. *Administrazioa Euskaraz* 87: 32-34.
- Gaminde, I., Aurrekoetxea, G., Etxebarria, A., Garay, U. & Romero, A. (2014). *Ahoskera Lantzeko Argibideak eta Jarduerak. Laguntzarako materiala: Teoria eta Praktika*. Bilbao: UPV/EHU.
- Garay, U. & Etxebarria, A. (2010). Irakasleen komunikazio-gaintasuna: hausnarketarako gaia. *Uztaro* 73: 39-48.
- Garay, U., Etxebarria, A., Gaminde, I., Romero, A. & Larrea, K. (2011). Emozioak etorkinen hizkuntzetan: estrategia linguistikoak eta euron pertzepzioa, in Ballano, I. (coord.), *Evaluar la competencia lingüística del alumnado inmigrante*, 137-150. Bilbao: Deustuko Unibertsitatea.
- Gil-Gómez de Liaño, B., & Pascual-Ezama, D. (2012). La metodología Delphi como técnica de estudio de la validez de contenido. *Anales de Psicología* 28(3): 1011-1020.
- Grau i Tarruell, M. & Vilà i Santasusana, M. (2009). La competencia prosódica y la comunicación no verbal, in Vilà i Santasusana, M. (coord.), *El discurso oral formal. Contenidos de aprendizaje y secuencias didácticas*, 89-100. Barcelona: Graó.
- Gustafson-Capková, S. (2011). Emotions in Speech: Tagset and Acoustic Correlates. Speech technology. Term paper in Speech Technology 1. Stockholm: Swedish National Graduate School of Language Technology (GSLT), Stockholm University.
- Hoekert, M., Kahn, R. S., Pijnenborg M. & Aleman A. (2007). Impaired recognition and expression of emotional prosody in schizophrenia: review and meta-analysis. *Schizophrenia research* 96: 135-145.
- Hsieh, H. F. & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9): 1277-1288.
- Hsieh, Hsiu-Fang, & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qual Health Res.* 15(9): 1277-88.
- Kendon, A. (1972). Some relationships between body motion and speech, in Seigman, A. & Pope, B. (Ed.), *Studies in Dyadic Communication*, 177-216. New York: Pergamon Press.
- Leentjens, A., Wielaert, S., van Harskamp, F. & Wilmink, F. (1998). Disturbances of affective prosody in patients with schizophrenia; a cross sectional study. *Journal of Neurology, Neurosurgery, and Psychiatry* 64: 375-378.
- Leon, S. & Rodriguez, M. (2008). Aprosodia and its treatment. *Perspectives on Neurophysiology and Neurogenic Speech and Language Disorders* 18(22): 66-72.
- Liscombe, J. (2007). *Prosody and speaker state: paralinguistics, pragmatics, and proficiency*. Columbia: Columbia University.
- Llisterri, J. (2016). *Los elementos suprasegmentales*. Universitat Autònoma de Barcelona. URL: [http://liceu.uab.es/~joaquim/phonetics/fon\\_prosod/suprasegmentales.html](http://liceu.uab.es/~joaquim/phonetics/fon_prosod/suprasegmentales.html) [14/03/18].
- Loehr, D. P. (2004) *Gesture and intonation*. Georgetown: Georgetown University.
- Martínez, A., Felizzola, C. A. & Matallana, D. L. (2015). Valoración de prosodia espontánea afectiva y análisis de discurso en pacientes con esquizofrenia y demencia frontotemporal (DFT) variante lingüística. *Revista colombiana de psiquiatría*, 44-1: 13-19.
- Massaro, D. W. (1989). Testing between the TRACE model and the fuzzy logical model of speech perception. *Cognitive Psychology* 21(3): 398--421
- Matsumoto, D. (2006). Are cultural differences in emotion regulation mediated by personality traits? *Journal of Cross-Cultural Psychology* 37: 421-437.
- McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.
- McQuown N. A. (1971). *The natural history of an interview*. Chicago: University of Chicago Library.
- Mozziconacci, S. (1998). *Speech Variability and Emotion: Production and Perception*. Eindhoven: Proefschrift.
- Murphy, D. & Cutting, J. (1990). Prosodic comprehension and expression in schizophrenia. *Journal of Neurology, Neurosurgery and Psychiatry* 53: 727-730.
- Navas, E., Hernández, I., Castelruiz, A., Sánchez, J. & Luengo, I. (2004). Acoustic Analysis of Emotional Speech in Standard Basque for Emotion Recognition, in Sanfeliu, A. et al. (Ed.), *CIARP 2004, LNCS 3287*, 386-393. Berlin: Springer-Verlag.
- Navas, E., Hernández, I., Luengo, I., Sánchez, J. & Saratxaga, I. (2005). Analysis of the Suitability of Common Corpora for Emotional Speech Modeling in Standard Basque, in Matousek, V. et al. (Ed.), *Lecture Notes on Artificial Intelligence*, vol. 3658: 265-272. Springer.
- Navas, E., Hernández, I., Luengo, I., Sainz, I., Saratxaga, I. & Sanchez, J. (2007). Meaningful Parameters in Emotion Characterisation, in Esposito, A. et al. (Ed.) *Verbal and Nonverbal Commun Behaviours. Lecture Notes on Computer Science*, vol. 4775: 74-84. Springer.
- Pereira, C. (2000). Dimensions of Emotional Meaning in Speech, in *ISCA workshop on Speech and Emotion*. Belfast.
- Pike, K. L. (1967). *Language in Relation to a Unified Theory of the Structure of Human Behaviour*. The Hague: Mouton.
- Ross, E. D. Orbelo D. M., Cartwright, J., Hansel, S., Burgard, M., Testa, J. A. & Buck, R. (2001). Affective-prosodic deficits in schizophrenia: profiles of patients with brain damage and comparison with relation to schizophrenic symptoms. *Journal of Neurology, Neurosurgery, and Psychiatry* 70: 597- 604.
- Saá, N. (2001). Lenguaje y hemisferio derecho. *Cuadernos de neurología*, XXV. URL: <http://publicacionesmedicina.uc.cl/cuadernos/2001/06.html> [13/03/18]

- Sainz, I., Saratxaga, I., Navas, E., Hernáez, I., Sanchez, J., Luengo, I. & Odriozola, I. (2008). Subjective Evaluation of an Emotional Speech Database for Basque, in *Proc. of the Sixth International Language Resources and Evaluation (LREC'08)*. Paper 437.
- Saratxaga I, Navas E., Hernaez I., & Luengo I. (2006). Designing and Recording an Emotional Speech Database for Corpus Based Synthesis in Basque, in *Proceedings of the LREC 2006*, 2126-2129. Genova.
- Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information* Vol 44(4): 695-729.
- Sidera, F., Serrano, J. & Amadó, A. (2012). El papel de la prosodia verbal y de la expresión facial en el entrenamiento de la falsa creencia a través de la denominación. *Boletín de Psicología* 105: 43-59.
- Soskin W. F. & Kauffman, P. E. (1961). Judgements of Emotions in Word-free Voice Samples. *Journal of Communication* 11(2): 73-80.
- Srinivasan, R. J. & Massaro, D. W. (2003). Perceiving from the face and voice: Distinguishing statements from echoic questions in English. *Language and Speech*, 46(1): 1-22.
- Van Rijn S., Aleman A., Swaab H. & Kahn R. S. (2005). Neurobiology of emotion and high risk for schizophrenia: role of the amygdala and the X-chromosome. *Neuroscience & Biobehavioral Reviews* 29: 385-397.