

# Expresión de emociones en la síntesis de voz en contextos narrativos

Virginia Francisco, Pablo Gervás, Raquel Hervás

Departamento de Sistemas Informáticos

Universidad Complutense de Madrid

Ciudad Universitaria

Madrid

[virginiafrancisco@gmail.com](mailto:virginiafrancisco@gmail.com) , [pgervas@sip.ucm.es](mailto:pgervas@sip.ucm.es) , [raquelhb@fdi.ucm.es](mailto:raquelhb@fdi.ucm.es)

## Resumen

Muchas de las aplicaciones que se plantean en la actualidad en inteligencia ambiental serían inviables sin un interfaz simple e intuitivo que facilite la interacción con el usuario. La comunicación a través de voz es una solución ideal para este problema. La voz generada por los conversores texto-voz actualmente suena artificial y ésta es la principal causa de rechazo por parte del público general. La aceptación de la inteligencia ambiental en la vida cotidiana de la gente pasa por conseguir superar ese rechazo. Para conseguir que el sintetizador aparente “vida” interesa generar voz con distintos estados anímicos. La generación de voz emocional pretende conseguir emociones suficientemente claras para que no haya confusión en el oyente. En este artículo se revisan las teorías existentes para definir una escala emocional, y se describe la elección de una escala concreta. Basado en esa elección, se lleva a cabo un análisis de las características acústicas de los distintos estados emocionales, para intentar posteriormente regenerar las mismas a través del sintetizador. Estas cuestiones se ponen a prueba experimentalmente explorando la viabilidad de modelar las cadencias propias de la narración de historias a través de los parámetros de control de un conversor texto-voz.

## 1. Estado del arte

Se puede decir que los estudios acerca de la generación de voz con emoción se remontan a Fairbanks [5] a finales de los años 30. Desde entonces han sido muchas las investigaciones y

los artículos publicados. Sin embargo, aun quedan por cubrir una gran variedad de aspectos relevantes.

En primer lugar hay que tener en cuenta que se trata de un área de investigación muy compleja. La complejidad comienza en el concepto mismo de emoción. No existe aun, hoy en día, una teoría aceptada generalmente de las emociones, en su lugar, existen múltiples teorías cada una con un interés distinto y difíciles de integrar.

### 1.1. Significado de la palabra emoción

Scherer [13] define las emociones como un mecanismo flexible de adaptación a un ambiente cambiante.

Un artículo de Cowie y Cornelius [4] introduce los siguientes términos y definiciones:

- *Emociones extremas*: Este término denota una emoción totalmente desarrollada, la cual típicamente es intensa e incorpora la mayoría de los aspectos que se consideran relevantes en el síndrome de la emoción.
- *Emociones subyacentes*: Denotan el tipo de colorante emocional que es parte de la mayoría o de todos los estados mentales. Un ejemplo es la cordialidad que a pesar de no ser una emoción por sí misma, implica una serie de emociones positivas.

### 1.2. Clasificación de las emociones

Para la investigación será necesario seleccionar un subconjunto de estas emociones que cumpla ciertos requisitos. Existen distintas teorías para realizar esta selección:

1. *Categorías emocionales*: Se trata de la técnica más comúnmente empleada, emplea palabras para denotar las emociones. Los lenguajes humanos proporcionan una serie de etiquetas que clasifican los distintos estados emocionales. Esta teoría engloba a otras tres:

- *Emociones básicas*: El número de emociones básicas suele ser pequeño (en los primeros estudios menos de 10, en los más recientes entre 10 y 20)
  - *Emociones súper ordinarias*: Existen una serie de categorías que son más fundamentales que otras en el sentido de que incluyen en sí mismas a las otras [11] [14]
  - *Emociones esenciales del día a día*: La teoría se ejemplifica en el trabajo de Cowie [4] que propone un diccionario de las emociones básicas inglesas. Comenzando con una lista de términos emocionales de la literatura se insta a los sujetos a seleccionar un subconjunto que represente de manera apropiada las emociones relevantes de la vida diaria. El resultado fue un subconjunto de 16 emociones.
2. *Descripciones basadas en psicología*: Según esta teoría el aspecto esencial de una emoción es el estado del cuerpo que tiene asociado [1][18]
3. *Descripciones basadas en la evaluación*: Estas teorías describen las emociones desde el punto de vista de las evaluaciones que implican [15]
4. *Dimensiones emocionales*: Esta alternativa considera las emociones como dimensiones. Las dimensiones emocionales son una representación simplificada de las propiedades esenciales de las emociones. Evaluación (positiva / negativa), y activación (activa / pasiva) son las dimensiones más importantes, en algunas ocasiones se complementan con la dimensión poder (dominante / sumiso) [20]

A la hora de generar voz con emoción el primer paso a tomar será decidir cuáles son las emociones que se desean modelar. Un buen comienzo quizá sería escoger un conjunto pequeño de emociones para, a continuación, completarlas hasta obtener un conjunto de emociones extenso. Para el primer paso sería bueno elegir cualquiera de los primeros métodos: emociones básicas, emociones súper ordinarias o las emociones del día a día. Cuando ya tengamos estas emociones bien modeladas habría que

ampliarlas incorporando emociones menos extremas, para esta tarea se podría elegir cualquiera de los otros métodos: descripciones basadas en psicología, basadas en evaluación o dimensiones emocionales. Otra forma de actuar sería comenzando desde un principio con cualquiera de las clasificaciones más extensas.

### 1.3. Fuentes de datos de voz con emoción

Numerosas investigaciones han desarrollado distintas estrategias para obtener una base de datos de voz con emoción. A continuación se describen los métodos más utilizados:

1. *Actores*: La técnica más antigua y más frecuentemente usada consiste en obtener los datos de grabaciones de actores.
2. *Lectura expresiva de material emocional*: Se trata de una variante al uso de actores propuesta por Nick Campbell [3] que propone tener lectores que lean textos con un contenido verbal apropiado con la emoción que se desea transmitir.
3. *Ocurrencias naturales*: Un estudio que trabaja con la generación espontánea de estados de ánimo es el dirigido por Scherer, Ladd y Silverman [16]. Grabaron entrevistas entre trabajadores sociales y actores que simulaban ser clientes. Tan solo se utilizaba el material grabado de los trabajadores sociales.
4. *Producción de emociones*: Se trata de inducir a las personas por medio de distintos métodos: hipnosis, imaginación, películas, música... a expresar una emoción.

Los métodos mostrados varían en cuanto al control que se puede ejercer sobre la señal del habla, se podría decir que de más a menos control se ordenarían de la siguiente manera: actores, lectura de textos emocionales, producción de emociones, observación del habla natural.

Cada uno de los métodos anteriormente mencionados encajan mejor o peor dependiendo del dominio del estudio que se este realizando:

- *Investigación de emociones extremas*: El método más adecuado será la utilización de actores.
- *Investigación de emociones subyacentes*: Lo mejor es observar el habla natural.
- *Estudios centrados en el hablante*: La mejor elección será la producción de emociones.

#### 1.4. Prosodia y emoción

Existen numerosos estudios para obtener las reglas prosódicas que intervienen en la generación de voz con emoción. En todos estos estudios los parámetros globales como son el nivel de la frecuencia fundamental, la escala de frecuencia fundamental y la velocidad del habla de la prosodia se tratan como universales, al menos siempre que el número de categorías emocionales sea pequeño.

Para la síntesis de voz las variables acústicas más interesantes son aquellas que se pueden controlar en los sistemas de síntesis de voz.

A la hora de generar un sistema capaz de emitir voz con emoción será necesario contar con una correspondencia entre las distintas emociones que se desean modelar y las características de cada uno de los parámetros que intervienen en la prosodia.

#### 1.5. Paradigmas de evaluación

A continuación se muestran los posibles paradigmas de evaluación:

1. *Elección forzada*: Este tipo de evaluación ha sido empleada en gran cantidad de estudios de generación de voz con emoción [2] [7] [8]. Consiste en facilitar a los sujetos que van a realizar la evaluación un conjunto finito de posibles respuestas que engloban todas las emociones que han sido modeladas. Las ventajas de este sistema son que es fácil de llevar a cabo, proporciona una medida simple de reconocimiento y permite comparar distintos estudios. Pero por otro lado tiene una desventaja y es que no proporciona información acerca de la calidad del estímulo desde el punto de vista de naturalidad y veracidad.
2. *Elección libre*: En este caso la respuesta no se restringe a un conjunto cerrado de emociones [9] [17]. Esta especialmente indicado para encontrar fenómenos no esperados durante el experimento.
3. *Elección libre modificada*: Murray y Arnot [9] y luego Stallo [19] introdujeron una serie de modificaciones al sistema anterior: introdujeron categorías de distracción, la categoría "otros", se emplearon textos neutrales junto con los textos con emoción y la diferencia entre el reconocimiento del texto neutro y del texto con

emoción se toma como medida del impacto de la prosodia en la percepción.

## 2. El cuenta-cuentos. Soluciones para la expresión de emociones

Los cuentos tratan de que los niños comprendan mejor sus sentimientos, por ello los cuentos deben tratar de resumir las emociones que experimentan la mayoría de los niños en su difícil camino a la madurez: alegría, tristeza, enfado, miedo, celos...

Cuando se cuenta un cuento se tiende a exagerar. La voz de la persona o personas que cuentan el cuento será un instrumento, tan importante como las palabras, para que el niño pueda inferir las emociones de los personajes. Podemos afirmar por tanto, que las emociones que se expresan en los cuentos son emociones extremas, no emociones subyacentes.

### 2.1. Emociones en el cuenta-cuentos

Las emociones en los cuentos tienen dos funciones fundamentales: transmitir al oyente la personalidad de sus personajes y transmitir al oyente las emociones que sienten cada uno de los personajes en todo momento.

La relación entre la cualidad de la voz del hablante y sus características psicológicas ha sido probada por muchos autores de forma experimental.

Para explicar los marcadores de la voz que van a dar "vida" y personalidad a nuestro cuento, nos vamos a basar en los estudios de K. R. Scherer [13]. Scherer define dos conceptos fundamentales acerca de la personalidad:

- *Caracteres de la personalidad*: Diferencias individuales reales y duraderas.
- *Atribuciones de personalidad*: Inferencias que los que observan hacen de los caracteres de la personalidad.

Una vez distinguidos ambos conceptos, podemos decir que mediante los marcadores de personalidad el hablante externaliza determinados caracteres y el oyente infiere y le atribuye esos mismos caracteres.

De entre las distintas clasificaciones de las emociones expresadas en el habla, quizá la que más se adecua a las necesidades de un cuenta-cuentos es la clasificación de emociones básicas,

ya que son este tipo de emociones las que se intentan transmitir a la hora de contar un cuento, además, cuando contamos un cuento tendemos a exagerar las emociones por lo que con un pequeño conjunto de emociones extremas tendremos suficiente. Para comenzar hemos seleccionado las siguientes emociones básicas: enfado, alegría, tristeza, miedo y sorpresa.

En los cuentos generados por nuestro cuenta-cuentos el único personaje que habla con su propia voz es el narrador, por lo que hemos decidido que las emociones aparecerán reflejadas en fragmentos del cuento. El narrador intentará imprimir una emoción distinta a cada frase, de modo que cuando este contando un pasaje triste del cuento su voz transmitirá tristeza y alegría cuando se trate de un pasaje alegre y feliz.

## 2.2. Marcación de los cuentos

Para facilitar la obtención de texto con emociones hemos adaptado un módulo ya existente de generación de texto en lenguaje natural en el campo de los cuentos fantásticos [6]. A partir de una entrada que representa las acciones que forman la trama del cuento y de información semántica sobre los personajes, lugares y objetos involucrados en las mismas, se genera el texto correspondiente en lenguaje natural.

La entrada principal del módulo de generación es el plan de trama. En él se da una secuencia de escenas, y dentro de cada una de estas escenas una lista de acciones. Cada acción tendrá como parámetros ciertos personajes, lugares u objetos, y equivaldrá a una frase concreta en el texto final.

En el proceso de generación de texto, ya sea para síntesis de voz como para su lectura, es necesario utilizar información semántica sobre los elementos involucrados en la historia, tanto personajes, lugares y objetos, como las acciones en las que se ven involucrados. En nuestro generador tenemos esta información semántica en una base de conocimiento [10] donde se almacenan, entre otros, la lista de personajes, lugares y objetos que pueden aparecer en los cuentos, sus atributos correspondientes y las relaciones que existen entre ellos.

Es también en esta base de conocimiento donde almacenamos la información necesaria para el marcado del texto con emociones. En nuestro trabajo consideramos que hay dos elementos que

se deben tener en cuenta a la hora de trabajar con emociones: los personajes y las acciones. Es obvio que en un cuento habrá ciertos personajes que irán acompañados de emociones alegres, como por ejemplo el caballero o la princesa, o de emociones de miedo, como la bruja o el dragón. Lo mismo ocurre con las acciones. Una acción triste podría ser la pérdida de un ser querido, o una acción sorprendente la aparición de un objeto mágico. Esta información irá en la base de conocimiento acompañando al personaje o a la acción correspondiente.

El marcado del texto en nuestro generador se realiza en la fase de lexicalización, donde para cada elemento del texto se escoge la etiqueta léxica correspondiente del vocabulario disponible. Es también en esta fase donde se elige el marcado para expresar las emociones deseadas.

En los primeros experimentos de marcado de cuentos para síntesis de voz hemos decidido marcar los textos a nivel de oraciones. Para decidir la emoción con la que marcar cada frase se tienen en cuenta tanto los personajes involucrados en la misma, como la acción en sí. De esa forma una acción alegre en la que está involucrado un personaje triste dará lugar a una oración con emoción neutra. Una acción alegre con un personaje alegre dará por supuesto una emoción alegre. Sin embargo, habrá emociones en acciones o personajes que deberán prevalecer sobre las demás. Por ejemplo, una acción sorprendente dará lugar a una emoción de sorpresa, independientemente de las emociones relacionadas con los personajes involucrados.

## 2.3. Fuentes de datos de voz con emoción

A la hora de determinar la relación entre los parámetros de la voz y las distintas emociones empleadas por nuestro cuenta-cuentos hemos recurrido a la grabación de cuentos leídos por actores. Esta elección se debe a que en el cuenta-cuentos como ya hemos mencionado anteriormente vamos a tratar emociones extremas y la utilización de actores es la mejor elección cuando lo que se busca es la investigación de emociones extremas.

A partir de los cuentos contados por actores hemos empleado la herramienta Praat [12] para analizarlos, con esta herramienta se pueden modificar distintos parámetros así como obtener

gráficos en los que se pueden observar perfectamente las variaciones de la voz y de los distintos parámetros de esta. Con Praat hemos obtenido la frecuencia fundamental media y su rango de variación así como la velocidad de lectura en las distintas emociones.

## 2.4. Prosodia

Los aspectos de la voz que se ha demostrado que actúan como marcadores de personalidad son: frecuencia fundamental, volumen, cualidad de la voz y fluidez. Para este trabajo hemos asumido que los aspectos de la voz necesarios para modelar las distintas emociones son: tono, volumen y fluidez. Es decir, nos olvidamos de la cualidad de la voz. No existe unanimidad a la pregunta: ¿Es fundamental la cualidad de la voz para generar voz con emoción? Nosotros hemos asumido que la cualidad de la voz no es fundamental a la hora de generar emociones.

Los parámetros de la prosodia que hemos variado para obtener las emociones se muestran en la siguiente tabla, las columnas representan los parámetros que hemos tomado como necesarios para modelar las distintas emociones y las filas identifican los valores que tomarán estos parámetros en la expresión de las emociones modeladas.

	Volumen	Velocidad (Pal/min.)	Pitch	Rango Pitch
Neutral	0.9	120	100	11
Enfado	1.0	145	100	30
Sorpresa	1.0	120	125	20
Alegría	1.0	155	135	14
Tristeza	0.8	110	90	7
Miedo	1.0	135	175	24

Tabla 1. Parámetros de la prosodia en las emociones

## 2.5. Evaluación

En la evaluación se han tratado de medir dos aspectos fundamentalmente:

- Grado de reconocimiento de cada una de las emociones.
- Influencia del significado en el reconocimiento de emociones.

Para conseguir estas medidas se han hecho una serie de distinciones que atienden a los textos que intervienen y el tipo de pruebas realizadas.

Los textos que han formado parte de las pruebas son de dos tipos:

- Cuentos con emociones en las distintas frases que lo componen. Se ha tomado uno de los cuentos generados, "Hansel y Gretel". Se ha seleccionado este cuento y no otro por tratarse de un cuento que reunía todas las emociones modeladas y por ser el cuento que más frases emocionantes contenía de entre todos los generados.
- La misma frase leída con cada una de las 5 emociones modeladas.

Mediante esta división se pretende medir el impacto del significado de las frases en la adjudicación de emociones, así como el impacto que tienen la consecución de emociones intercaladas con texto neutro.

Se han realizado dos tipos de pruebas:

- Pruebas de elección forzada: Se da a elegir entre las 5 emociones modeladas y los usuarios deben seleccionar una de estas 5 emociones.
- Pruebas de elección libre: A la lista anterior se han añadido 3 emociones más: culpa, vergüenza y envidia, y la categoría "Otras".

Mediante esta división se pretende medir por un lado la capacidad de reconocimiento de las emociones modeladas así como la aparición de confusiones con emociones no modeladas.

Para las pruebas se han tomado a 10 personas, primero se les han realizado las pruebas de elección libre sobre un cuento y sobre la misma frase repetida 5 veces cada una con una de las 5 emociones modeladas, a continuación se han realizado las pruebas de elección forzada. Mientras escuchaban el cuento iban asignando a todas aquellas frases "emocionantes" la emoción que se correspondía con lo que estaban percibiendo. En el caso de las frases, los evaluadores marcaban cada una de ellas con una emoción de las posibles.

Los resultados obtenidos se muestran en las siguientes tablas, las filas representan las emociones que se pretendían modelar (E: Enfado, S: Sorpresa, A: Alegría, T: Tristeza, M: Miedo, N: Neutral, V: Vergüenza, C: Culpa, EN: Envidia), y las columnas representan las emociones percibidas por los evaluadores.

	E	S	A	T	M	V	C	EN	O	N
Enfado	90%	0%	0%	0%	0%	0%	10%	0%	0%	0%
Sorpresa	23%	52%	5%	0%	12,5%	0%	2,5%	5%	0%	0%
Alegría	25%	10%	55%	0%	0%	0%	0%	0%	10%	0%
Tristeza	0%	0%	0%	87,5%	0%	10%	2,5%	0%	0%	0%
Miedo	35%	15%	0%	0%	40%	0%	0%	10%	0%	0%
Neutral	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%

Tabla 2. Cuento con elección libre

	E	S	A	T	M	V	C	EN	O	N
Enfado	60%	0%	30%	0%	0%	0%	0%	10%	0%	0%
Sorpresa	0%	10%	10%	20%	10%	30%	20%	0%	0%	0%
Alegría	20%	20%	10%	0%	20%	10%	10%	10%	0%	0%
Tristeza	0%	0%	0%	70%	0%	0%	10%	20%	0%	0%
Miedo	0%	0%	0%	0%	50%	10%	20%	20%	0%	0%
Neutral	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%

Tabla 3. Frases con elección libre

	Enfado	Sorpresa	Alegría	Tristeza	Miedo	Neutral
Enfado	90%	10%	0%	0%	0%	0%
Sorpresa	20%	70%	5%	0%	5%	0%
Alegría	20%	0%	80%	0%	0%	0%
Tristeza	0%	0%	0%	100%	0%	0%
Miedo	25%	35%	0%	0%	40%	0%
Neutral	0%	0%	0%	0%	0%	100%

Tabla 4. Cuento con elección forzada

	Enfado	Sorpresa	Alegría	Tristeza	Miedo	Neutral
Enfado	70%	0%	30%	0%	0%	0%
Sorpresa	0%	50%	20%	0%	30%	0%
Alegría	10%	30%	50%	0%	10%	0%
Tristeza	0%	0%	0%	100%	0%	0%
Miedo	20%	20%	0%	0%	60%	0%
Neutral	0%	0%	0%	0%	0%	100%

Tabla 5. Frases con elección forzada

La primera afirmación que podemos realizar a la vista de los resultados es que las emociones siempre se diferencian del texto neutro, no ha habido confusión en este sentido.

También podemos afirmar que existe una pequeña diferencia entre los resultados de los cuentos y los de las frases, los aciertos en los cuentos son más elevados que en las frases lo

que nos lleva a concluir que el significado influye de manera positiva en la asociación de una emoción a unas determinadas características de la voz. En cuanto al reconocimiento de cada una de las emociones los resultados que se extraen de las pruebas anteriores son los siguientes:

- **Enfado:** El enfado es una de las emociones más reconocidas, de entre todas las emociones modeladas tan solo se confunde, en un porcentaje significativo, con la alegría. A la hora de evaluar el resto de emociones, todas ellas, excepto la tristeza, se confunden en algún momento de las pruebas con el enfado. Esto quizás es debido a que la voz neutra que se ha elegido como base es una voz muy seria, con una frecuencia fundamental baja, lo que conlleva que cualquier emoción tenga una percepción de seriedad.
- **Sorpresas:** Se trata de una de las emociones menos logradas, en los experimentos realizados tan solo obtiene un porcentaje de acierto elevado en el experimento de elección forzada con el cuento, es decir, en el reconocimiento de la sorpresa intervienen tanto el significado como la existencia de un conjunto restringido de emociones.
- **Alegría:** La alegría es una de las emociones menos reconocidas, obteniendo un reconocimiento especialmente bajo en el experimento de elección libre con las frases, esto es debido a que cuando existen muchas posibilidades y el significado no ayuda a determinar la emoción, en el caso de las frases alegres se realiza la asignación de emoción de manera aleatoria, ya que se confunde con todas las emociones ofrecidas como posibles.
- **Tristeza:** Sin duda alguna se trata de la emoción más lograda, obteniendo un 100% de reconocimiento en los experimentos de elección forzada.
- **Miedo:** La media de acierto en este caso ronda en todos los experimentos el 50%, las confusiones más importantes se dan con el enfado y la sorpresa.

### 3. Conclusiones

Una vez realizadas todas las pruebas llegamos a la conclusión de que habría que estudiar más a fondo emociones como la sorpresa y la alegría ya que no son fácilmente identificables. En cuanto al enfado, la tristeza y el miedo se puede decir que están bastante logradas y que con unos ligeros retoques podrían llegar a ser reconocidas sin ninguna duda. Si bien es cierto, que obtener un 100% de reconocimiento es prácticamente imposible, ya que el reconocimiento de emociones es algo subjetivo, y bastante difícil. A la hora de realizar las pruebas los evaluadores han dejado clara la dificultad que tenemos las personas a la hora de determinar las emociones que percibimos.

Los siguientes pasos a llevar a cabo en la generación de voz con emociones en el cuento pasan por realizar un análisis más exhaustivo de los valores de los parámetros de la voz que intervienen en las emociones. La asignación de las emociones en los cuentos tomados como base para el análisis debería ser realizada por un conjunto más extenso de personas, de modo que esta asignación sea más general y menos subjetiva, quizá sería bueno que estas mismas personas fuesen en la última fase las que realizaran la evaluación. Otro punto importante a tener en cuenta en un trabajo futuro sería el de buscar más voces de base, tanto de mujeres como de hombres, y con ellas llegar a unas emociones más reconocibles, sobre todo en lo que a la alegría y la sorpresa se refiere. Una vez conseguido este objetivo, la siguiente meta será el modelado de emociones subyacentes.

### Referencias

- [1] Alter, K., Rank, E., Kotz, S. A., Toepel, U., Besson, M., Schirmer, A., and Friederici, A. D. (2000). Accentuation and emotions – two different systems? In Proceedings of the ISCA Workshop on Speech and Emotion, pages 138–142, Northern Ireland. <http://www.qub.ac.uk/en/isca/proceedings>
- [2] Cahn, J. E. (1990). The generation of affect in synthesized speech. Journal of the American Voice I/O Society, 8:1-19.

- [3] Campbell, N. (2000). Databases of emotional speech. In Proceedings of the ISCA Workshop on Speech and Emotion, pages 34-38, Northern Ireland. <http://www.qub.ac.uk/en/isca/proceedings>
- [4] Cowie, R. and Cornelius, R.R. (2003) Describing the emotional states that are expressed in speech. *Speech Communication Special Issue on Speech and Emotion*, 40(1-2): 5-32.
- [5] Fairbanks, G. and Pronovost, W. (1939). An experimental study of the pitch characteristics of the voice during the expression of emotion. *Speech Monograph*, 6:87-104.
- [6] Gervás, P., Díaz-Agudo, Belén, Peinado, F. and Hervás, R. (2004) Story plot generation based on CBR. In Anne Macintosh, Richard Ellis, and Tony Allen, editors, 12<sup>th</sup> Conference on Applications and Innovations in Intelligent Systems, Cambridge, UK. Springer, WICS series.
- [7] Montero, J.M. (2003) Estrategias para la mejora de la naturalidad y la incorporación de variedad emocional a la conversión texto a voz en castellano. Tesis Doctoral, ETSI Telecomunicación, UPD, Director: J. M. Pardo.
- [8] Mozziconacci, S. J. L. (1998). *Speech Variability and Emotion: Production and Perception*. PhD thesis, Technical University Eindhoven.
- [9] Murray, I. R. and Arnott, J. L. (1995). Implementation and testing of a system for producing emotion-by-rule in synthetic speech. *Speech Communication*, 16:369-390.
- [10] Peinado, F., Gervás, P., Díaz-Agudo, B.: "A Description Logic Ontology for Fairy Tale Generation". In Veale T., Cardoso A., Camara Pereira, F., Gervás, P. (Eds.): 4th International Conference on Language Resources and Evaluation, Procs. of the Workshop on Language Resources for Linguistic Creativity, LREC'04, 56-61. 29th May, Lisbon, Portugal. LREA, 2004.
- [11] Ortony, A., Clore, G.L., and Collins, A. (1988). *The Cognitive Structure of Emotion*. Cambridge University Press, Cambridge, UK.
- [12] <http://www.fon.hum.uva.nl/praat/>
- [13] Scherer, K. R. (1979). *Personality markers in speech* by K. R. Scherer and H. Giles (eds): Social markers in speech. Cambridge: Cambridge University Press.
- [14] Scherer, K. R. (1984). Emotion as a multicomponent process: A model and some cross-cultural data. *Review of Personality and Social Psychology*, 5:37-63.
- [15] Scherer, K. R. (1984). On the nature and function of emotion: A component process approach. In Scherer, K. R. and Ekman P., editors, *Approaches to emotion*, pages 293-317. Erlbaum, Hillsdale, NJ.
- [16] Scherer, K. R., Ladd, D. R., and Silverman, K. (1984). Vocal cues to speaker affect: Testing two models. *Journal of the Acoustic Society of America*, 76(5):1346-1356.
- [17] Schröder, M. (1999). Can emotions be synthesized without controlling voice quality? *Phonus 4*, Research Report of the Institute of Phonetics, University of the Saarland, pages 37-55. <http://www.dfki.de/~schroed>
- [18] Smith, C.A. (1989) Dimensions of appraisal and physiological response in emotion. *Journal of Personality and Social Psychology*, 56(3):339-353.
- [19] Stallo, J. (2000). *Simulating emotional speech for a talking head*. Honour's thesis, School of Computing, Curtin University of Technology, Australia. <http://www.computing.edu.au/~stalloj/projects/honours>
- [20] Wundt, W. (1896). *Grundriss der Pshychologie*. Verlag von Wihelm Engelmann, Leipzig.