

# Mediación emocional en sistemas de Comunicación Aumentativa y Alternativa

Nestor Garay, Julio Abascal, Luis Gardezabal

Laboratorio de Interacción Persona-Computador para Necesidades Especiales  
Informatika Fakultatea. Euskal Herriko Unibertsitatea  
Manuel Lardizabal 1; E-20018 Donostia  
E-mail: {nestor, julio, luisg}@si.ehu.es  
Teléfono: +34 943018000; Fax: +34 943219306

## Resumen

Muchas personas que carecen de la capacidad de hablar debido a diversos tipos de discapacidades motoras y del lenguaje utilizan dispositivos informáticos de ayuda a la comunicación. Estos dispositivos aportan un medio de comunicación alternativa muy útil pero, en la mayoría de los casos, lento y carente de expresividad. En los últimos años se ha dedicado un gran esfuerzo a la mejora de la velocidad de comunicación mediante diversos métodos, tales como la predicción de palabras o el diseño de dispositivos de entrada adaptables, obteniéndose mejoras interesantes. Sin embargo, el problema de la falta de expresividad en la comunicación no ha sido suficientemente estudiado. El reciente auge de la *computación emocional* abre una vía a la inclusión de información de contexto que permita reflejar los sentimientos, estados de ánimo y actitudes, en las ayudas a la comunicación. El presente artículo hace un estudio de la aplicabilidad de la computación emocional a las ayudas a la comunicación y describe un prototipo que está siendo usado para evaluar la calidad de la comunicación lograda.

**Palabras clave:** Mediación emocional, Interfaces para personas con discapacidad, Modelado de la conversación, Comunicación Aumentativa y Alternativa

## 1. Introducción

Las personas con severas discapacidades motoras y del habla<sup>1</sup> necesitan recurrir a sistemas de Comunicación Alternativa<sup>2</sup> para hacerse entender

<sup>1</sup> Esto incluye a usuarios afectados por un variado espectro de discapacidades, con diferentes orígenes y niveles de afectación: parálisis cerebral, apraxia, esclerosis múltiple y esclerosis lateral amiotrófica, afasia, etc. [AAC89].

<sup>2</sup> Se denomina Comunicación Alternativa a cualquier forma de comunicación distinta del habla e incluye todos los métodos de comunicación utilizados por las personas que, por diferentes causas (discapacidad, enfermedad, envejecimiento, accidente, etc.) carecen de lenguaje oral [Basil88]. La Comunicación Aumentativa tiene el doble objetivo de promover y apoyar el habla y garantizar una

por los demás. Frecuentemente, la comunicación alternativa requiere dispositivos externos al usuario, en cuyo caso se habla de comunicación asistida. Con este fin, y desde hace tiempo, se vienen utilizando diversos tipos de dispositivos de ayuda a la comunicación, tales como los tableros silábicos por ejemplo. Sin embargo, la difusión de los ordenadores personales ha hecho posible el diseño de ayudas a la comunicación informatizadas, frecuentemente llamadas comunicadores, que toman un papel de intermediario activo que puede dinamizar enormemente la comunicación.

---

forma de comunicación alternativa [Tetzchner 96].

Estos dispositivos de comunicación no son más que ordenadores portátiles dotados de una interfaz especial. Dadas las restricciones motoras de los usuarios, es necesario sustituir los dispositivos estándar de entrada/salida por otros adecuados a las características físicas de los usuarios. Por ejemplo, se puede sustituir el teclado por un sistema de barrido controlado por un único pulsador, accionable con cualquier parte del cuerpo sobre la que se conserve algún control residual. También es posible usar algunas técnicas de interacción avanzadas que han experimentado progresos interesantes, tales como el seguimiento del movimiento del ojo (eye tracking), [Jacob95], o la captación de algún tipo de señal eléctrica cerebral<sup>3</sup>, para controlar el movimiento del ratón sobre la pantalla) [Lusted96, Rohde98].

En todos los casos, el usuario tiene que deletrear o silabear las palabras que quiere decir. El ordenador las reconstruye, si es necesario, las corrige y las lee en alto mediante un sintetizador de voz. Para mejorar la velocidad de producción de palabras, se han diseñado métodos de predicción que aprovechando la redundancia y la estructura gramatical del idioma permiten adelantar la palabra que el usuario está tratando de escribir. De todas maneras, en el mejor de los casos la velocidad de comunicación que se consigue no es equivalente a la del lenguaje hablado<sup>4</sup>. Esto crea problemas de integración a los usuarios de dispositivos de comunicación alternativa ya que difícilmente pueden participar en una conversación normal.

El uso de dispositivos como intermediarios en la conversación presenta otro problema. El comunicador transmite mensajes en un estilo neutro, carentes del contexto que usualmente acompaña a la comunicación oral (gestos expresiones, tono de voz...) y que da importante información del estado de ánimo y, en general, de la situación personal del hablante. De hecho, el contexto no sólo permite transmitir información expresiva, sino que también ayuda a interpretar adecuadamente el mensaje, eliminando la ambigüedad en frases con doble sentido, expresiones irónicas o sarcásticas, etc. Por lo tanto, el usuario de sistemas de comunicación

---

<sup>3</sup> Señales eléctricas tales como las ondas cerebrales (electroencefalogram, EEG), potenciales evocados (evoked potentials, EP), las producidas por el movimiento de los ojos (electroocular potentials, EOG) o los potenciales musculares medidos en el cráneo o la cara (muscle potentials EMG).

<sup>4</sup> Según Alm et al., en una conversación normal en inglés se dicen por término medio unas 120-180 palabras por minuto, mientras que la composición de mensajes por medio de dispositivos de comunicación permite velocidades de 2 a 10 palabras por minuto [Alm92].

asistida frecuentemente se ve en dificultades para expresar los matices que son usuales en la conversación humana.

El problema de mejorar la velocidad de comunicación que permite el dispositivo intermediario ha sido ampliamente tratado mediante diversas técnicas, tales como la optimización de la disposición del conjunto de selección, el uso de teclados reducidos con codificación ambigua, la adaptación automática de la velocidad de barrido, [Gardeazabal99], los sistemas de anticipación, el modelado de la conversación, etc. [Garay00]. El segundo problema –el de añadir expresividad a la comunicación alternativa desarrollada a través de dispositivos intermediarios– se le ha dedicado, en general, menor atención. Sin embargo, merecen ser mencionados los trabajos que viene realizando para el inglés el Departamento de Computación Aplicada de la Universidad de Dundee en Escocia<sup>5</sup> [Murray91, Murray00, Alm93, Alm99] y, para el castellano, el Grupo de Tecnología del Habla de la Universidad Politécnica de Madrid<sup>6</sup> [Montero98].

En este artículo se presenta el trabajo realizado en el Laboratorio de Interacción Persona-Computador para Necesidades Especiales de la Universidad del País Vasco-Euskal Herriko Unibertsitatea para mejorar la expresividad de los dispositivos de comunicación asistida mediante la aplicación de técnicas de computación emocional.

## 2. Comunicación emocional

La interacción humana incluye información sobre los afectos, emociones, estados de ánimo, etc., de los interlocutores, que se transmite de manera explícita a través del lenguaje y de manera implícita a través de la comunicación no verbal [Knapp80]. Como se ha mencionado anteriormente, la información no oral, que frecuentemente se transmite mediante gestos, actitudes, expresiones, etc., es de gran importancia en la comunicación humana, ya que tiene un gran efecto sobre la disposición comunicativa de los interlocutores y sobre la inteligibilidad del discurso. Sin embargo, y de acuerdo con [Picard97, Picard98], se ha apreciado que estas características que se asocian a las relaciones interpersonales también aparecen en la comunicación con los ordenadores. Incluso aunque el usuario no espera que el ordenador sea

---

<sup>5</sup>Department of Applied Computing, University of Dundee, Scotland. <http://www.computing.dundee.ac.uk>

<sup>6</sup>Departamento de Ingeniería Electrónica-E.T.S.I. Telecomunicación-Universidad Politécnica de Madrid. <http://www-gth.die.upm.es/> y <http://www.die.upm.es/>

capaz de procesar información emocional, ésta tiende a aparecer espontáneamente cuando la interfaz de usuario lo permite (por ejemplo, en las interfaces con entrada por voz). La información así transmitida puede ser muy relevante a la hora de determinar características importantes de la interacción persona-ordenador, como son los intereses, actitudes y fines del usuario. Por esta razón el principal objetivo de la computación emocional es captar y procesar la información afectiva con el fin de mejorar la comunicación entre la persona y el ordenador.

Por tanto, la hipótesis básica de la computación emocional es que si se tienen en cuenta las características afectivo-emocionales del usuario se puede mejorar notablemente la comunicación. Pero, ¿cómo puede detectar el ordenador el estado de ánimo del usuario? Para esta tarea se puede emplear el análisis de los gestos, la detección de diversos parámetros de la voz (volumen, tono...), la fuerza e insistencia con que se teclea o utiliza el ratón e, incluso, la detección de parámetros físicos corporales tales como el pulso, la conductividad de la piel y otra serie de parámetros que pueden variar según el estado de ánimo. En algunos casos, la detección requiere métodos invasivos (hace falta aplicar algún tipo de sensor o vestir al usuario con alguna prenda especial), aunque la miniaturización que se está consiguiendo en los circuitos permitirá que su tamaño se reduzca de modo que puedan ser integrados en pendientes, gafas, etc. [Picard98].

Avanzando en esa misma línea, se supone que la información que el ordenador transmite al usuario será mejor entendida y aceptada por éste si contiene un contexto de información emocional similar al que se produce en la interacción humana. Por lo tanto sería interesante que, por su parte, el ordenador transmitiera su "propia" información emocional. Por ejemplo, su "estado de ánimo", mediante la modulación de la voz sintética o mediante gestos realizados por personajes o avatares que lo representen. Esto permitiría no sólo transmitir información textual, sino también la expresividad con que ésta se ofrece.

## 2.1 Mediación emocional

Dentro de la computación emocional, se denomina mediación emocional al caso concreto en el que se hace uso de un ordenador como intermediario entre la comunicación de varias personas, reflejando el ánimo que presentan los interlocutores [Picard97]. De este modo se trata de compensar el filtrado de información emocional que realizan los dispositivos de comunicación asistida debido a la pérdida de la

información no verbal. Tal como se ha manifestado previamente, la mediación emocional tiene aplicación directa dentro de la Comunicación Aumentativa y Alternativa, que es el objeto de este artículo. También se conocen aplicaciones a otros tipos de comunicaciones mediadas, tales como las telecomunicaciones textuales (correo electrónico "emocional", chats "emocionales", etc.).

En el siguiente apartado se presenta un sistema para de ayuda a la comunicación por teléfono de personas con discapacidad motora y oral, para posteriormente describir su ampliación mediante técnicas de mediación emocional.

## 3. El sistema GESTELE

El punto de partida es un sistema de ayuda a la comunicación, denominado GESTELE, desarrollado previamente en el Laboratorio de Interacción Persona-Computador para Necesidades Especiales (ver figura 1), con el objetivo de que las personas con grandes restricciones en sus movimientos y que carecen de habla, puedan mantener comunicaciones telefónicas [Garay95, Garay00, Juanena01, Pérez99].

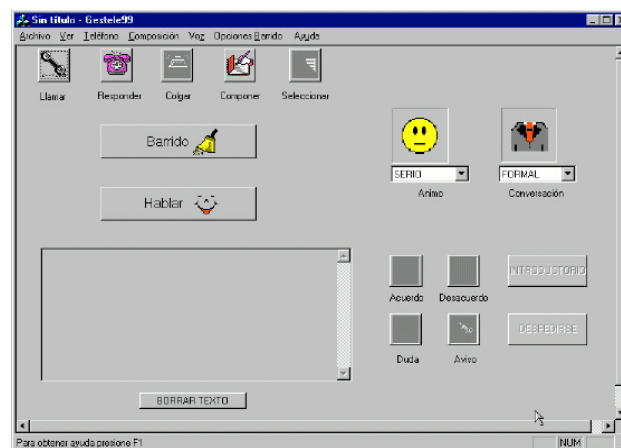


Figura 1. Interfaz de GESTELE99.

GESTELE permite la composición de frases mediante un sistema de barrido controlado por un único pulsador. Para mejorar la velocidad de comunicación cuenta con un predictor/corrector de palabras. Con el objetivo de mejorar la velocidad de comunicación y minimizar la aparición de tiempos muertos mientras el usuario compone una nueva palabra, el sistema dispone de frases prealmacenadas que son utilizadas automáticamente por el sistema dependiendo del estado en el que esté la conversación. Una vez introducida o seleccionada

una frase, puede leerla en voz alta mediante un sistema comercial de síntesis de voz conectado al auricular del teléfono.

Dispone de un modelo de la conversación similar al empleado por el sistema CHAT, [Alm92], que permite representar internamente el estado de la conversación por medio de un autómata [Garay95, Garay00]. Por ejemplo, en el estado inicial, el sistema selecciona aleatoriamente un saludo entre todos los disponibles. Posteriormente podrá emitir frases de compromiso o de introducción al tema, tras las cuales puede pasar al tema central y, finalmente, concluir con frases de despedida. Aunque el sistema es capaz de decir frases previamente almacenadas en función del estado de la conversación, el usuario controla la evolución del diálogo mediante las diversas opciones presentes en la interfaz.

La interfaz de usuario consta de una serie de botones que permiten seleccionar fácilmente mediante barrido o selección directa los controles más usuales del sistema (ver figura 1). Hay botones para controlar el teléfono: llamar, responder o colgar; para editar: componer o seleccionar mensajes; para sintetizar textos por medio de un programa de traducción texto-a-voz (“Hablar”) o para su reproducción (“Borrar texto”). Además existen opciones que permiten incluir directamente frases de relleno habituales que muestran acuerdo, desacuerdo o dudas. Se pueden mandar avisos para indicar que el usuario está componiendo un mensaje adecuado a la conversación en curso, de modo que el interlocutor no interprete el silencio como un corte en la comunicación. Otras opciones permiten introducir el tema principal de la conversación o despedirse; y, finalmente, activar o desactivar el barrido de las opciones. En función de algunas de estas opciones el autómata puede cambiar de estado siguiendo la evolución de la conversación

El resto de los controles se utilizan para el desarrollo de la conversación, pero no cambian el estado del autómata. Entre ellos está el editor de frases. Hay que reseñar que, asociados al editor de frases, se incorporan las utilidades de mostrar un teclado virtual en la pantalla para ir escribiendo mensajes e incluso una ventana asociada a un predictor de palabras que muestra las palabras más frecuentes con el comienzo que haya escrito el usuario, con el objetivo de minimizar el número de pulsaciones requerido para componer los mensajes [Juanena01] (ver figura 2).

## 4. Emocionalidad en GESTELE

La interfaz del sistema GESTELE ha mejorado algunos de sus aspectos relativos a la comunicación emocional. Por ejemplo, se cuenta con botones para que el usuario refleje su estado de ánimo así como el tipo de conversación que pretende llevar a cabo<sup>7</sup>. Para cada uno de los casos, se contemplan cuatro posibles valores. Los estados de ánimo tratados son: alegre, triste, enfadado y serio, mientras que los tipos de conversación contemplados son: formal, informal, jocoso y agresivo<sup>8</sup>. Esta información se tiene en cuenta para la generación de frases automáticas. Así pues, dependiendo de estos valores, el conjunto de frases que se ofrece, así como el modo en que se expresan en voz alta, será diferente. Para ello, la base de datos de frases las clasifica según el estado de la conversación a que correspondan (por ejemplo, saludos frente a despedidas), y les asigna valores relativos al estado de ánimo y el tipo de conversación para el que sean más apropiadas (como por ejemplo, alegre y jocoso o enfadado y agresivo).

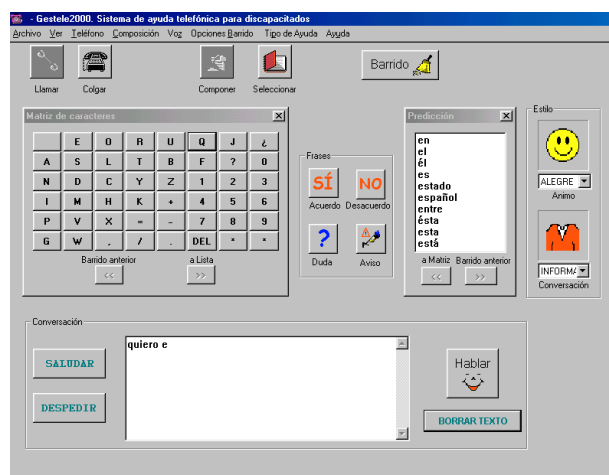


Figura 2. Interfaz de la versión actual de GESTELE.

### 4.1 Transmisión de características emocionales

Para reflejar el estado de ánimo del usuario se acondicionan el tono, el volumen y la velocidad de

<sup>7</sup>Mientras que el estado de ánimo depende básicamente del usuario que utiliza el sistema, el tipo de conversación depende tanto del usuario como del interlocutor con que se esté comunicando y del propio desarrollo de la conversación.

<sup>8</sup>Aunque parezca extraña la inclusión de estados y estilos tales como enfadado y agresivo, una de las necesidades de los usuarios de sistemas de mediación emocional es poder expresar estados de ánimo que normalmente están excluidos de los dispositivos de comunicación políticamente correctos [Alm99].

la voz emitida por el sistema (de manera análoga a la reflejada en [Iriondo00].) que son parámetros ajustables en el sintetizador de voz empleado. Para ello, se ha asociado una tríada (tono, volumen, velocidad) a cada estado de ánimo y estilo de conversación, que es modificable por el usuario. También puede seleccionar entre un conjunto de voces predeterminadas el tipo de voz que mejor le representa y modificar algunas de las características de la misma para adaptarlas a su gusto.

En una primera aproximación, los valores del ánimo y el tipo de la conversación son fijos, a menos que el usuario decida cambiarlos, para lo que puede escoger una de entre las opciones presentes en las listas desplegables asociadas a dichas variables. De esta forma, el usuario mantiene en todo momento el control sobre estos valores que presenta el sistema. Sin embargo, y en consonancia con [Pérez99], las personas con dificultad motora y oral que intentan teclear lo más rápido posible suelen evitar modificar los parámetros afectivos debido al esfuerzo y al consumo de tiempo que ello les supone. Por ello, y con el fin de mejorar la velocidad de comunicación, se ha desarrollado un sistema que trata de inferir automáticamente el ánimo del usuario.

#### **4.2 Detección automática del estado de ánimo y del tono de la conversación**

Un sencillo modelo de usuario permite almacenar su estado de ánimo actual y modificarlo basándose en la observación de la evolución de la conversación. Concretamente, se realiza un análisis de los textos que va produciendo el usuario para detectar la aparición de una serie de palabras clave que permiten hacer hipótesis sobre su estado de ánimo. Para esto se ha compilado un diccionario de términos claves asociados a ánimos particulares y dotados con pesos relativos según lo ajustados que estén a cada tipo de ánimo. Este diccionario se ha construido de forma manual en función del léxico utilizado por el usuario.

#### **4.3 Evaluación**

A partir de evaluaciones con sujetos sin discapacidad se han obtenido unos resultados iniciales interesantes que han servido para mejorar el tratamiento automático de estado de ánimo y el tipo de conversación. Sin embargo es necesario refinar la asignación de parámetros de voz a cada estado de ánimo y a cada estilo de conversación, ya que son muy dependientes de la persona y su contexto cultural. En la actualidad se están llevando a cabo evaluaciones por medio de interacciones con usuarios finales para determinar la validez del sistema así como su usabilidad.

## **5. Conclusiones**

Tal como se ha visto la aplicación de la computación emocional a la Comunicación Aumentativa y Alternativa para mejorar la comunicación, rehabilitación e integración de personas con discapacidad resulta muy prometedora. Concretamente, el trabajo realizado demuestra que la mediación emocional puede servir para mejorar la transmisión de información de contexto en los dispositivos de ayuda a la comunicación mejorando las posibilidades expresivas del usuario y la interpretabilidad del mensaje.

El sistema que se ha diseñado añade información del estado de ánimo y del estilo de conversación a un dispositivo de comunicación asistida para conversiones telefónicas basado en el modelado de la conversación, que cuenta con síntesis de voz. Las frases producidas tratan de añadir contenido emocional mediante la modificación de parámetros básicos de la voz sintetizada. En una primera aproximación, el propio usuario selecciona los valores correspondientes. Sin embargo, para evitar el esfuerzo añadido que esto requiere se ha diseñado un modelo de usuario con información emocional que se actualiza automáticamente en función del léxico utilizado por el usuario.

## **Agradecimientos**

Los autores agradecen al resto de los miembros del Laboratorio de Interacción Persona-Computador para Necesidades Especiales su colaboración en el desarrollo de GESTELE.

## **Referencias**

- [AAC89] Monografía sobre tipos de usuarios. *Alternative and Augmentative Communication*, vol. 5, no. 1. (1989).
- [Alm92] N. Alm, J. L. Arnott and A. F. Newell: 'Prediction and Conversational Momentum in an Augmentative Communication System'. *Communications of the ACM*, vol. 35, no. 5. pp 46-57 (1992).
- [Alm93] N. Alm, I. R. Murray, J. L. Arnott, and A. F. Newell. 'Pragmatics and affect in a communication system for non-speakers'. *Journal of the American Voice Society*, vol.13. pp.1-15. (1993).
- [Alm99] Alm, N., Ellis, K. 'Computer-assisted swearing'. *Communicating Together*, vol. 16 no. 2. pp 4-7. (1999).
- [Basil88] C. Basil y R. Puig de la Bellacasa. *Comunicación aumentativa*. INSERSO, (1988).

- [Garay95] N. Garay-Vitoria, J. G-Abascal, and S. Urigoitia-Bengoa. 'Application of the Human Conversation Modelling in a Telephonic Aid'. *Proceedings of the 15th International Symposium on Human Factors in Telecommunications (HFT '95)*, Melbourne (Australia). Pp. 131-138. (1995).
- [Garay00] N. Garay. 'Sistemas de Predicción Lingüística. Aplicación a Idiomas con Alto y Bajo Grado de Flexión, en el Ámbito de la Comunicación Aumentativa y Alternativa'. Tesis Doctoral. Facultad de Informática. UPV/EHU. San Sebastián, (2000).
- [Gardezabal99] L. Gardezabal. 'Aplicaciones de la Tecnología de Computadores a la Mejora de la Velocidad de Comunicación en Sistemas de Comunicación Aumentativa y Alternativa'. Tesis Doctoral. Facultad de Informática de la Universidad del País Vasco (UPV/EHU). San Sebastián, (1999).
- [Iriondo00] I. Iriondo, R. Guaus, A. Rodríguez, P. Lázaro, N. Montoya, J. M<sup>a</sup> Blanco, D. Bernadas, J. M. Oliver, D. Tena, L. Longhi. 'Validation of an Acoustical Modelling of Emotional Expression in Spanish Using Speech Synthesis Techniques'. *Proceedings of the ISCA Workshop on Speech and Emotion*. <http://www.qub.ac.uk/en/isca/index.htm>. (2000).
- [Jacob95] R. J. K. Jacob, 'Eye Tracking in Advanced Interface Design'. In *Virtual Environments and Advanced Interface Design*, ed. by W. Barfield and T.A. Furness, pp. 258-288, Oxford University Press, (1995).
- [Juanena01] A. Juanena. 'Diseño e Implementación de una Ayuda Técnica para Discapacitados'. Proyecto Fin de Carrera. Facultad de Informática. UPV/EHU, (2001).
- [Knapp80] M. L. Knapp. 'Essentials of nonverbal communication'. Holt, Rinehart & Winston, (1980).
- [Lusted96] H. S. Lusted and R. B. Knapp. 'Controlling computers with neural signals'. *Scientific American*, vol. 275, no. 4, pp. 82-87, (1996).
- [Montero98] J. M. Montero, J. M. Gutiérrez, S. Palazuelos, E. Enríquez, S. Aguilera, J. M. Pardo. 'Emotional Speech Synthesis: From Speech Database to TTS'. *Proceedings of the 5th International Conference on Spoken Language Processing*. Sydney, Australia, (1998).
- [Murray00] R. Murray, (Emotion in concatenated speech", in: *Proceedings of the IEE Seminar State of the Art in Speech Synthesis*. London. pp.7/1-7/6. (2000).
- [Murray91] I. R. Murray, J. L. Arnott, N. Alm and A. F. Newell. 'A Communication System for the Disabled With Emotional Synthetic Speech Produced by Rule'. *Proc. Eurospeech '91*, pp. 311-314. (91)
- [Murray96] R. Murray, J. L. Arnott and E. A. Rohwer, 'Emotional stress in synthetic speech: Progress and future directions', *Speech Communication* vol. 20, no. 1-2, pp. 85-91, (1996)
- [Pérez99] S. Pérez. *Diseño e Implementación de una Ayuda Técnica para Discapacitados*. Proyecto fin de Carrera. Facultad de Informática. UPV/EHU, (1999).
- [Picard97] R. W. Picard. 'Affective Computing'. MIT Press, (1997).
- [Picard98] R. W. Picard. 'Towards Agents that Recognize Emotion'. *Acts Proceedings IMAGINA (Monaco)*, pp 153-165, (1998).
- [Rohde98] M. M. Rohde, S. L. BeMent, S. P. Levine, J. E. Huggins, R. K. Kushwaha. 'Signal to Noise Ratio Based Sorting of Voluntary Event Related Potential Averages for Assistive Technology Applications' *Proc of the RESNA 98 Annual Conf*. pp. 381-383, (1998).
- [Tetzchner93] S. V. Tetzchner & H. Martinsen 'Introducción a la enseñanza de signos y al uso de ayudas técnicas para a comunicación' (Adaptado por C. Basil). Visor, Madrid, (1993).