

# Accesibilidad para Discapacitados a través de Teléfonos y Servicios Móviles Adaptables

Diego López de Ipiña, Amparo Rodríguez e Iñaki Vázquez  
Facultad de Ingeniería  
Universidad de Deusto (Bilbao)  
{ dipina, amrodrig, ivazquez }@eside.deusto.es

## Resumen

El teléfono móvil es una herramienta que facilita nuestras vidas, nos permite contactar con cualquiera, entretenernos, guardar información donde y cuando queremos. Sin embargo, los teléfonos móviles no pueden o son difíciles de utilizar todavía por personas con deficiencias audiovisuales, motrices o de avanzada edad. Todo ello, a pesar de ofrecer cada día mejores capacidades multimedia y de comunicación (tanto de corto como largo alcance) y ser cada vez más fácilmente programables. Hasta el momento, las aplicaciones a los que estos avances en tecnología móvil han sido destinados son MMS (Mensajería Multimedia), y la reproducción de videos y música. En este artículo describimos un par de servicios de infraestructura que añaden accesibilidad a los teléfonos móviles para así facilitar la vida de aquellas personas que más lo necesitan, las personas con discapacidades físicas. A través de los servicios MobileSense y PhoneProxy, un móvil convencional se adapta a las necesidades de accesibilidad demandadas por personas invidentes, sordas o con problemas motrices.

## 1. Introducción

Los teléfonos móviles de última generación están adoptando paulatinamente capacidades hasta ahora sólo disponibles por ordenadores personales y PDAs (Personal Digital Assistant). Sus cada día mejores capacidades multimedia (captura de imágenes y reproducción de sonidos polifónicos), de comunicación (corto alcance mediante Bluetooth y largo alcance vía GPRS y UMTS) y, sobre todo, mejor capacidad de computo, nos permiten imaginar escenarios hasta hace poco impensables. Todas estas mejoras en capacidades técnicas están aderezadas con mecanismos cada vez más sencillos, y a su vez, poderosos para programar aplicaciones en dispositivos móviles. En este sentido tres plataformas de desarrollo (en el orden dado) dominan el mercado: MIDP 2.0 [1], Symbian [2], y SmartPhone OS [3].

La accesibilidad a los sistemas informáticos ha experimentado un progreso significativo en los últimos años. Sistemas operativos como Windows XP o Linux (tanto para el entorno gráfico Gnome como KDE) vienen ya equipados de serie con facilidades para la accesibilidad [4][5][6]. Algunos ejemplos de las utilidades provistas son:

- Magnificadores de pantalla y sintetizadores de texto a voz para personas con problemas visuales.
- Herramientas que transforman sonidos del sistema en mensajes visuales para sordos.
- Mejoras de acceso a teclado y ratón para personas con problemas motrices.

Por tanto, existe una plétora de utilidades tanto propietarias como de código abierto que permiten a las personas con discapacidades hacer uso de una herramienta tan útil en el mundo actual como es el ordenador, y por extensión Internet.

Consiguientemente, parece el momento oportuno para que los campos de accesibilidad software y tecnología móvil converjan. Tecnológicamente estamos en situación de integrar un sintetizador de texto a voz (TTS) en un teléfono móvil para asistir a invidentes, o somos capaces de llevar a cabo una videoconferencia tanto a través de GPRS como UMTS, para permitir a dos sordos comunicarse mediante gestos. Asimismo, la capacidad de comunicación vía Bluetooth [7] e Infrared [8] de los teléfonos móviles permite considerar escenarios en los que personas inválidas o con problemas motrices se sirvan de sus móviles como mandos a distancia universales que controlen componentes domóticos tanto en sus casas como en entornos de trabajo u ocio. En definitiva, se hace imperativo trasladar el progreso alcanzado en accesibilidad dentro del campo de los ordenadores personales a los teléfonos móviles.

Comenzaremos este artículo examinando el estado del arte en la implantación de accesibilidad en los teléfonos móviles. A continuación, analizaremos el estado del arte de los mecanismos que permiten añadir accesibilidad a las aplicaciones informáticas: OCR, TTS y reconocimiento de voz. Como consecuencia de estos dos estudios resaltaremos el largo camino a recorrer que todavía resta, y la inexistencia de herramientas libres que adapten los móviles, de manera genérica, a las necesidades de las personas con discapacidades. Para resolver esta carencia, definiremos un par de servicios móviles, libremente accesibles (en el sentido económico), que añaden propiedades de accesibilidad a los teléfonos móviles de personas con discapacidades.

## **2. Estado del Arte de la Accesibilidad en los Teléfonos Móviles**

Hasta el momento la industria ha concentrado su mayor esfuerzo en transformar a los móviles en dispositivos más accesibles para el colectivo de invidentes. Dos soluciones software: Talks y Mobile Accessibility, y una solución hardware, Owasys, destacan sobremedida en este terreno.

Talks [9] añade a un móvil Symbian de la Serie 60 o 90 la capacidad del habla. El principal cometido de esta aplicación es que una persona invidente sepa en todo momento sobre qué elemento de la interfaz del móvil se encuentra y cuáles son las acciones disponibles en el contexto actual. Talks utiliza el sintetizador de voz ETI Eloquence [10].

Mobile Accessibility [11] de la empresa Code Factory cambia la interfaz del teléfono para que personas con deficiencias visuales pero no ciegas puedan todavía manejar el teléfono. A su vez, también usa un sintetizador de voz, proporcionado por SVOX [12], para indicar a un usuario invidente el contexto actual dentro del sistema de menús del teléfono. Mientras que Talks mantiene los menús por defecto de un móvil, Mobile Accessibility simplifica las opciones de menú disponibles en el móvil, para así reducir su complejidad.

La empresa OWASYS [13] ha creado móviles adaptados tanto a gente invidente como a personas mayores. El modelo 22C es un móvil sin pantalla que viene acompañado de un sintetizador de voz proporcionado por Babel [14], y provee las funcionalidades típicas de cualquier móvil convencional, como mensajería de texto, lista de contactos y telefonía. El modelo 112C es un teléfono simplificado a la máxima expresión, también carece de pantalla, pero dispone de un botón de emergencias especiales. Es un móvil que intenta emular al máximo el comportamiento de un teléfono convencional de línea fija (da tono de llamada, permite preasignar teclas a números importantes). Está orientado al colectivo

de gente de avanzada edad, que necesita valerse por si misma. Otras empresas han propuesto soluciones similares como Mobile Senior [15] o Secufone [16].

En cuanto a las personas con deficiencias de oído, las adaptaciones que se han realizado sobre móviles son prácticamente nulas. La gente sorda se ha servido principalmente de los mensajes de texto para comunicarse y del vibrador del móvil para darse cuenta cuándo se reciben mensajes de entrada. La única manera de permitir que una persona no sorda se comunique por teléfono con otra sorda es a través de la intermediación de un estenógrafo que en tiempo real escribe lo que oye para que la persona sorda pueda leer lo que dice el otro interlocutor. La organización benéfica británica RNID para la ayuda a la gente sorda ofrece en asociación con BT un servicio que responde a estas características y que es denominado TypeTalk [17]. A su vez, en España la Secretaría de Estado de Servicios Sociales, Familias y Discapacidad ofrece un servicio similar de intermediación [18]. El principal inconveniente de estos sistemas es su requerimiento de la intermediación de un operador humano. Idealmente sería más conveniente que la labor de traducción de voz a texto y, viceversa, fuera realizada automáticamente por un ordenador.

Finalmente, para las personas con discapacidades motrices han comenzado a aparecer sistemas de control domótico que permiten el uso del móvil como un mando de control remoto. Usando la capacidad de comunicación Bluetooth o Infrared de los teléfonos móviles es posible a través de pasarelas Bluetooth/Infrared a estándares domóticos como X10, controlar los dispositivos domóticos de un espacio. Un buen ejemplo de este tipo de sistemas es el Sistema Domótico Integral ideado por Code Factory [19].

### **3. Estado del Arte en TTS, Reconocimiento de Voz, OCR y MIDP**

De la revisión del estado del arte sobre accesibilidad en teléfonos móviles se deduce que la incorporación de las siguientes capacidades en teléfonos móviles, o, en su defecto, en servicios móviles (pasarelas) que representan a los móviles, se hace capital:

- TTS (Text-To-Speech), es decir, la capacidad de transformar texto en voz.
- Reconocimiento de Voz (Speech Recognition), permitir que sin previo entrenamiento la dicción de un interlocutor se transforme en texto.
- OCR (Optical Character Recognition), o transformación del contenido textual de imágenes escaneadas en cadenas de caracteres.
- APIs para el acceso programático a los mecanismos de captura y reproducción de sonidos e imágenes, así como de comunicación Bluetooth/Infrared, dentro de un móvil.

En este artículo nos concentraremos en explorar las herramientas libres disponibles en estas materias. Así, podremos determinar la infraestructura libre ya existente en la que basar los servicios de accesibilidad para móviles que pretendemos.

La transformación de texto a voz (TTS) es una técnica madura para la cual existen múltiples herramientas capaces de generar voz en tiempo real. Festival [20] y FreeTTS [21] son buenos ejemplos de tecnologías libres que realizan TTS en ordenadores personales. Por su parte, Flite [22] es un motor TTS libre para PDAs que usan Linux. Desafortunadamente, no existen motores TTS libres que puedan ejecutarse en teléfonos Symbian o MIDP 2.0, similares a los sistemas propietarios mencionados, Eloquence y SVOX.

El reconocimiento de voz continuo sin entrenamiento e independiente del interlocutor no es todavía lo suficientemente robusto cuando consideramos vocabularios de miles de palabras [23]. Los productos IBM ViaVoice [24] y Dragon NaturallySpeaking [25] son

los líderes del mercado en este sector. ViaVoice ofrece una implementación para dispositivos empotrados denominada, Embedded ViaVoice. Estos productos son muy fiables bajo condiciones controladas: bajo sonido de ambiente, reconocimiento dependiente de una sola persona, y velocidad de lectura en vez de conversación. En el plano de herramientas libres, Sphinx4 [26] representa una opción equiparable. De hecho, Sphinx4 está diseñado para permitir el reconocimiento de voz independientemente del interlocutor.

El reconocimiento óptico de caracteres (OCR) es un campo muy maduro para el que existen múltiples soluciones libres y propietarias con un funcionamiento general muy robusto. Como solución de código abierto destacamos GOOCR [27].

Los mayoría de los teléfonos de última generación, tales como el Nokia 6600 o Sony Ericsson P900, son programables por medios de las APIs Java de J2ME [1] o C++ de Symbian [2]. J2ME se está convirtiendo en el estándar *de facto* para la programación de dispositivos móviles ya que es soportada por casi todos los últimos modelos. El perfil J2ME más utilizado es el Mobile Information Device Profile (MIDP). Todos los teléfonos compatibles Java en el mercado soportan MIDP 1.0 o superior. MIDP 1.0 ofrece paquetes para la conexión por GPRS, creación de interfaces de usuario, o para guardar datos de manera persistente. La última versión de MIDP, 2.0, provee funcionalidad extra que puede ser muy útil en la tarea de añadir accesibilidad a los móviles: conectividad por red segura con HTTPS, aplicaciones push o comunicación con el puerto serie.

Existen dos extensiones MIDP 2.0 opcionales que facilitan nuestra objetivo de añadir mayor accesibilidad a los teléfonos móviles: la Mobile Media API (MMAPI), y la Bluetooth API. MMAPI es un paquete opcional que permite la reproducción y captura de audio y video. La Bluetooth API añade capacidad de programación Bluetooth a los dispositivos móviles.

A pesar de la popularidad de MIDP, productos como Eloquence o SVOX que permiten TTS en teléfonos móviles han utilizado las APIs de Symbian. Aunque hay menos teléfonos que soportan Symbian que los que soportan Java, Symbian permite acceder a la parte interna del móvil e implementar algoritmos que requieren un alto rendimiento de ejecución.

En conclusión, dado el estado del arte actual parece complicado poder ver pronto teléfonos móviles con capacidad de reconocimiento de voz continuo. Sin embargo, es muy factible encontrar pronto móviles que soporten OCR, vía su cámara, o TTS. Nosotros pretendemos contribuir, en un futuro próximo, con versiones libres de TTS y OCR ejecutables en un móvil. De hecho ya hemos contribuido con un software que permite a un móvil leer códigos de barras [28]. Actualmente estamos extendiendo este sistema para que pueda leer cadenas de caracteres. En el resto de este artículo nos concentramos en discutir el diseño de un par de servicios de accesibilidad destinados a usuarios invidentes o sordos de teléfonos móviles. El desarrollo de estos dos servicios, MobileSense y PhoneProxy, fue motivado por la carencia detectada de aplicaciones y servicios de accesibilidad de carácter libre para teléfonos móviles.

#### 4. La Plataforma MobileSense

La plataforma MobileSense persigue los siguientes objetivos:

1. Estar al alcance de la mayoría de los dispositivos móviles de última generación. Es decir, estar basada en tecnología MIDP 2.0.
2. Ser una plataforma de accesibilidad universal genérica, o lo que es lo mismo no estar restringida a un colectivo de personas con una discapacidad particular.

3. Poder funcionar tanto en ambientes controlables (casa, trabajo) como no controlables (calle o campo). Es decir, poder delegar el procesamiento de imágenes o sonido, o la síntesis de voz a servidores locales vía Bluetooth o remotos vía UMTS o GPRS.

La Figura 1 muestra la arquitectura software cliente/servidor de MobileSense. La parte cliente la conforma una aplicación MIDP 2.0 instalable en cada móvil. Esta aplicación permite al usuario la captura de imágenes y sonido que son enviadas al servidor de MobileSense para su procesamiento. Las respuestas enviadas por el servidor a la aplicación, en forma de voz o subtítulos, son reproducidas por el player de audio y subtítulos integrados en esta aplicación cliente.

La parte servidora de la aplicación MobileSense procesa las imágenes o sonido capturado y enviado por los clientes. El servidor MobileSense puede procesar dos tipos de imágenes: aquellas correspondientes a fragmentos de texto en los que se aplica OCR por medio de GOOCR o aquellas que contienen los códigos de barras circulares del sistema TRIP [28]. Además, el servidor de MobileSense es capaz de efectuar reconocimiento de voz por medio de Sphinx4 y sintetizar voz a partir de texto por medio de FreeTTS. El proceso de reconocimiento de voz genera subtítulos que son reproducidos en la parte cliente de MobileSense, mientras que FreeTTS genera ficheros .wav que son reproducidos en el cliente.

La transmisión de contenido multimedia en ambos sentidos se realiza por medio de Bluetooth cuando los clientes y el servidor están cercanos, a una distancia máxima de 100 metros. Cuando el servidor de MobileSense no está al alcance del cliente, éste envía las peticiones de procesamiento por medio de GPRS o UMTS.

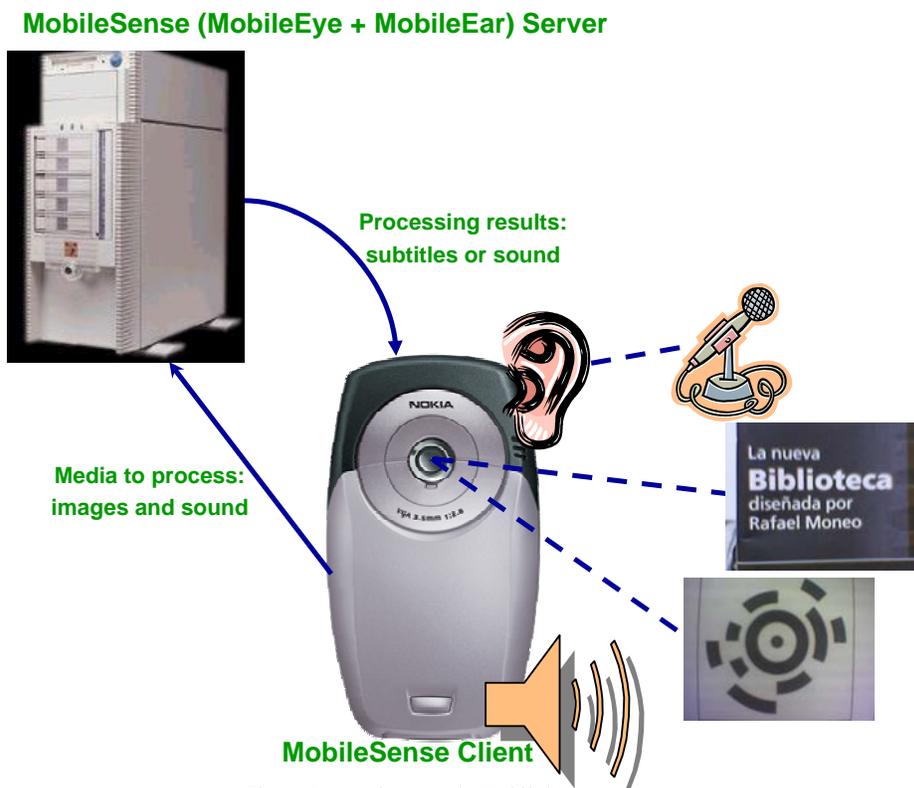


Figura 1. Arquitectura de MobileSense

## 5. La Plataforma PhoneProxy

La segunda solución que presentamos en nuestro intento de transformar los teléfonos móviles en dispositivos más accesibles a las personas con deficiencias auditivas o visuales es PhoneProxy. Esta solución persigue un enfoque similar al de TypeTalk [17] y el Centro de Intermediación [18] pero, en este caso, orientado a la gente invidente. La idea es permitir que cualquier invidente pueda acceder a la funcionalidad de un teléfono móvil (mensajería de texto, acceso a contactos o llamar) por medio de un servicio móvil que hace uso de la tecnología VoiceXML [29].

VoiceXML permite a los usuarios navegar por Internet desde el teléfono. Mediante VoiceXML es posible diseñar un tipo de páginas web que le hablan al usuario e interactúan con él mediante una interfaz vocal, sustituyendo el componente visual, que ha sido tradicionalmente el único disponible en la web. La tecnología VoiceXML es muy similar a los sistemas automatizados de atención al cliente y no requieren por parte del usuario más que un teléfono fijo o móvil convencional. En el pasado, aplicando VoiceXML hemos creado aplicaciones como DirectTraffic, un portal multidispositivo de información de tráfico, que permitía el acceso mediante voz a la información en tiempo real de incidencias de tráfico (cortes de carreteras, mejor ruta, etc..) desde el teléfono móvil.

En PhoneProxy, el usuario simplemente marca el número de teléfono del servicio donde un “contestador automático”, gobernado por VoiceXML, le indica la serie de opciones disponibles:

1. Efectuar llamada
2. Buscar persona en contactos
3. Leer mensajes de texto recibidos
4. Enviar mensajes de texto

El invidente por medio de su voz puede controlar la interfaz del PhoneProxy y seleccionar la acción deseada.

En la implementación de este servicio nos hemos servido de la plataforma de VoiceXML suministrada por Loquendo [30]. Por medio de VoiceXML creamos la interfaz de voz que representa a la aplicación PhoneProxy. La parte dinámica de este portal web de voz utiliza una base de datos en el servidor de PhoneProxy que permite guardar los detalles sobre todos los usuarios del sistema, incluyendo sus contactos y mensajes recibidos. El servidor de PhoneProxy tiene acoplado al puerto serie (RS-232) un módem GPRS a través del cual la aplicación recibe y envía mensajes y llamadas. PhoneProxy impone el requisito de que las personas que deseen enviar mensajes de texto a un usuario de PhoneProxy lo tengan que hacer al número teléfono del servidor, precediendo el mensaje real con un código de identificación del usuario. PhoneProxy utiliza el software Sphinx4 para transformar mensajes vocales en textuales. Para ello, combina tecnología de Voz con T9 [31] (Text on 9 keys) para facilitar la ardua tarea del reconocimiento de voz independientemente de quién hable.

## 6. Conclusión

Este trabajo ha descrito dos servicios que añaden propiedades de accesibilidad a los teléfonos móviles: MobileSense y PhoneProxy.

La plataforma MobileSense está basada en tecnologías libres para OCR, TTS y reconocimiento de voz. La integración del software FreeTTS y Sphinx4 en MobileSense ofrece una solución para que personas con deficiencias auditivas puedan entablar

conversaciones telefónicas. La integración de GOCR con FreeTTS dentro de MobileSense permite a una persona invidente “ver” anuncios y letreros a través de su móvil.

Por otro lado, el servicio PhoneProxy proporciona, de manera remota, la funcionalidad de un teléfono móvil a gente invidente. Por medio del estándar VoiceXML permitimos al usuario interactuar con una pieza de software que actúa como si se tratase de su móvil. Con la ayuda de este software el usuario puede escuchar los mensajes de texto recibidos, dictar nuevos mensajes de texto, seleccionar un contacto y efectuar llamadas.

En esencia, a través de los servicios MobileSense y PhoneProxy hemos demostrado que el teléfono móvil puede ser considerado como un dispositivo sensorial que ayuda a la gente con discapacidades a “ver” y “escuchar”, y a poderse comunicar con otras personas de una manera natural.

Como línea futura de trabajo estamos considerando el desarrollo de una aplicación de videoconferencia que permita a dos personas sordas comunicarse usando el lenguaje de signos. Por otro lado, a través de nuestro proyecto EMI<sup>2</sup> [32] estamos trabajando en el uso del móvil como un agente software personal que nos sugiere acciones y permite controlar de manera cómoda el entorno que nos rodea. Una extensión interesante a PhoneProxy sería poder permitir que las llamadas gestionadas por este servicio pudiesen en ocasiones ser efectuadas vía voz sobre IP (reduciendo costes) [33].

## Agradecimientos

Los autores quieren agradecer la ayuda recibida de la Cátedra de Telefónica Móviles en la Universidad de Deusto (<http://www.ctme.deusto.es>) para la realización de este trabajo.

## Referencias

- [1] *J2ME Mobile Information Device Profile (MIDP)*, Sun Microsystems, <http://java.sun.com/products/midp/>, 2004
- [2] *Symbian OS – the mobile operating system*, Symbian Corporation, <http://www.symbian.com/>, 2004
- [3] *Windows Mobile-based Smartphones*, Microsoft Corporation, <http://www.microsoft.com/windowsmobile/smartphone/default.aspx>, 2004
- [4] *Windows XP Accessibility Resources*, Microsoft Corporation, Windows XP Accessibility Resources, <http://www.microsoft.com/enable/products/windowsxp/default.aspx>, 2004
- [5] *Gnopenicus*, <http://www.baum.ro/gnopenicus.html>, 2004
- [6] *The KDE Accessibility Project*, <http://accessibility.kde.org/>, 2004
- [7] *Bluetooth*. <http://www.bluetooth.org/>.
- [8] *Infrared Data Association*, <http://www.irda.org/>, 2004
- [9] *TALKS for Series and Nokia Communicator*, [http://www.talx.de/index\\_e.shtml](http://www.talx.de/index_e.shtml), 2004
- [10] *ETI Eloquence SF*, <http://www.scansoft.com/etisf/>, ScanSoft, 2004
- [11] *Mobile Accessibility*, <http://mobileaccessibility.codfact.com/>, 2004
- [12] *SVOX-TTS Engine*, <http://www.svox.com/site/pro448697/svo392819/svo556982697.asp?osLang=1>
- [13] *OWASYS 22C y 112C*, [http://www.owasys.com/index\\_en.php](http://www.owasys.com/index_en.php), 2004
- [14] *Babel Text-To-Speech*, <http://www.babeltech.com/Products.php?s=76&m=75>, 2004
- [15] *Mobile Senior*, Code Factory, <http://msenior.codfact.com/>, 2004
- [16] *Secufone*, <http://www.secufone.com/>, 2004
- [17] *RNID TypeTalk BT*, <http://www.rnid-typetalk.org.uk/frame.html>, 2004

- [18] Centro de Intermediación Telefónica para Personas Sordas o con Discapacidad Auditiva y/o de Fonación, [http://www.seg-social.es/imserso/discapacidad/dis\\_cenintel.html](http://www.seg-social.es/imserso/discapacidad/dis_cenintel.html), 2004
- [19] *Sistema Domótico Integral*, [http://www.codfact.com/cfflash/esp/index\\_esp.html](http://www.codfact.com/cfflash/esp/index_esp.html) (sección Domótica), Code Factory, 2004
- [20] *The Festival Speech Synthesis System*, <http://www.cstr.ed.ac.uk/projects/festival/>, The Centre for Speech Technology Research, University of Edinburgh , 2004
- [21] *FreeTTS 1.2beta2 - A speech synthesizer written entirely in the JavaTM programming language*, <http://freetts.sourceforge.net/docs/index.php>, 2004
- [22] *Flite: a small, fast run time synthesis engine*, <http://www.speech.cs.cmu.edu/flite/>, Speech Software at CMU, 2004
- [23] *Audio Visual Speech Recognition*, <http://www.research.ibm.com/AVSTG/srec.html>, IBM Research, 2004
- [24] *IBM ViaVoice*, <http://www-306.ibm.com/software/voice/viavoice/>, 2004
- [25] *Dragon NaturallySpeaking 7*, <http://www.scansoft.com/naturallyspeaking/>, 2004
- [26] Walker w., Lamere P., Kwok P., Raj B., Singh R., Gouvea E., Wolf P. y Woelfel J., *Sphinx-4: A Flexible Open Source Framework for Speech Recognition*, SMLI TR2004-0811, <http://cmusphinx.sourceforge.net/sphinx4/doc/Sphinx4Whitepaper.pdf>, 2004
- [27] Schulenburg J., *GOOCR: Open Source Character Recognition*, <http://jocr.sourceforge.net/index.html>, 2004
- [28] López de Ipiña D., *MobileEye: a Ringcode Reader for Mobile Phones*, <http://paginaspersonales.deusto.es/dipina/tripweb/services.htm>, Universidad de Deusto, 2004
- [29] *VoiceXML Forum*, <http://www.voicexml.org/>, 2004
- [30] *Loquendo Vocal Technology and Services*, <http://www.loquendocafe.com/index.asp>, 2004
- [31] *T9*, <http://www.t9.com/>, 2004
- [32] López de Ipiña D. y Vázquez I., *Servicios Móviles Sensibles al Contexto*, Ponencia presentada en Feria MovilForum 2003, Madrid, <http://paginaspersonales.deusto.es/dipina/publications/MobileEye.ppt>, Diciembre 2003
- [33] *Skype, Internet Telephony that just Works*, <http://www.skype.org/>, 2004