

## Las tecnologías del habla en lengua catalana

Joaquim Llisterri

Grup de Fonètica, Seminari de Filologia i Informàtica, Departament de  
Filologia Espanyola, Universitat Autònoma de Barcelona

La lengua catalana dispone en la actualidad de un conjunto de aplicaciones, prototipos y productos en el campo de las tecnologías del habla, fruto de los esfuerzos de grupos de investigación universitarios, de empresas del sector y de iniciativas institucionales. Este trabajo pretende ofrecer una panorámica, necesariamente no exhaustiva, de los trabajos recientemente llevados a cabo en un ámbito que, como reconoce, por ejemplo, el *III Pla de Recerca de Catalunya* (2001-2004), es de una importancia fundamental tanto para favorecer el acceso a la Sociedad de la Información y del Conocimiento a todos los ciudadanos como para la plena normalización de la lengua.

El desarrollo de estas tecnologías se ha llevado a cabo, predominantemente, en el marco de las universidades catalanas: el *Grup de Tractament de la Parla* del Departamento de Teoría de la Señal y Comunicaciones de la *Universitat Politècnica de Catalunya* (UPC), -integrante del centro TALP, *Centre de Tecnologia i Aplicacions del Llenguatge i de la Parla*, junto con el *Grup de Processat del Llenguatge Natural* del Departamento de Lenguajes y Sistema Informáticos de la UPC- el *Grup de Tractament de Veu* de la *Escola Universitària Politècnica de Mataró* (EUPM), y el *Àrea de Tecnologies de la Parla* de la sección de Teoría de la Señal del Departamento de Comunicaciones y Teoría de la Señal de *Enginyeria La Salle* en la *Universitat Ramon Llull* (URL) orientan sus actividades desde la perspectiva de la ingeniería de telecomunicaciones, mientras que grupos como el *Grup de Fonètica* del Seminario de Filología e Informàtica del Departamento de Filología Española de la *Universitat Autònoma de Barcelona* (UAB) se centran en el conocimiento fonético y lingüístico que se incorpora a las tecnologías del habla<sup>9</sup>.

Las tecnologías del habla en catalán se han desarrollado especialmente, en lo que a la empresa se refiere, en Telefónica I+D (Madrid), con el fin de ofrecer servicios telefónicos automáticos en esta lengua. Más recientemente, empresas como ATLAS, *Applied Technologies on Language and Speech*, estrechamente

---

<sup>9</sup> *Grup de Tractament de la Parla*: <http://gps-tsc.upc.es/veu/>; TALP: <http://www.talp.upc.es>; *Grup de Tractament de Veu*: <http://www.eupmt.es/imesd/telematica/veu/>; *Àrea de Tecnologies de la Parla*: [http://www.salleurl.edu/Eng/elsDCTS/tsenyal/tsenyal\\_areaparla.html](http://www.salleurl.edu/Eng/elsDCTS/tsenyal/tsenyal_areaparla.html); *Grup de Fonètica*: <http://liceu.uab.es>

vinculada al grupo de la UPC, ofrecen soluciones tecnológicas para la creación de aplicaciones que incorporen el tratamiento del habla<sup>10</sup>.

En el ámbito de las iniciativas institucionales cabe mencionar el CREL, *Centre de Referència en Enginyeria Lingüística*, creado en 1996 como parte de la red de Centros de Referencia en Investigación y Desarrollo de la *Generalitat de Catalunya*. Este centro, cuya finalidad es la coordinación e integración de los trabajos de diversos grupos universitarios<sup>11</sup> en torno a un conjunto de proyectos comunes, tiene como objetivo la creación de recursos, herramientas y prototipos que favorezcan el desarrollo de la ingeniería lingüística en catalán.

Tradicionalmente las tecnologías del habla se dividen entre las que proporcionan una salida vocal –síntesis-, las que permiten la entrada de información -reconocimiento- y las que facilitan la interacción con un sistema informático -sistemas de diálogo-. Entre las primeras, la aplicación más destacada es tal vez la conversión de texto en habla, que hace posible la transformación automática de un texto escrito en su equivalente sonoro.

Para el catalán, existen al menos tres sistemas de conversión de texto en habla, desarrollados respectivamente por la *Universitat Politècnica de Catalunya*, la *Universitat Ramon Llull* y por Telefónica I+D. El conversor de texto a habla de la UPC, comercializado a través de ATLAS, ha contado con el apoyo del CREL y, en este marco, con la participación de la UAB. Se trata de un sistema que actualmente se presenta en dos versiones -una convencional basada en difonemas y otra más avanzada utilizando técnicas de selección de unidades a partir de un corpus-, para el que se dispone de voz masculina y femenina; el sistema permite también, a través de la incorporación de marcas en el texto de entrada, intercalar enunciados sintetizados en castellano (Bonafonte *et al.* 1997, 1999)<sup>12</sup>.

El desarrollo de este conversor ha conllevado el diseño de herramientas como SEGRE, un transcriptor fonético automático para el catalán que cubre las cuatro principales variantes geográficas, utiliza SAMPA (*SAM Phonetic Alphabet*) para representar la transcripción<sup>13</sup> y se caracteriza por un entorno que permite

<sup>10</sup> Telefónica I+D: <http://www.tid.es/trabajo/nuestrosproductos/telefonía/home.html>; ATLAS: <http://www.atlas-cti.com>

<sup>11</sup> *Equip del Diccionari del Català Contemporani* y *Equip de les Oficines Lexicogràfiques* del *Institut d'Estudis Catalans*, *Grup de Gramàtica Teòrica* y *Seminari de Filologia i Informàtica* de la *Universitat Autònoma de Barcelona*, *Grup de recerca de Variació en el Llenguatge* de la *Universitat de Barcelona*, *Grup de Lingüística Aplicada* de la *Universitat de Girona*, *Grup de Processat del Llenguatge Natural* y *Grup de Tractament de la Parla* de la *Universitat Politècnica de Catalunya* y *Grup de recerca en Lingüística Aplicada al Tractament del Llenguatge* de la *Universitat Pompeu Fabra*. La entidad gestora del centro es el *Institut d'Estudis Catalans*. Más información sobre el CREL se encuentra en <http://www.crel.net>

<sup>12</sup> Puede escucharse una demostración en <http://gps-tsc.upc.es/veu/soft/demos/tts.php3> y en <http://www.atlas-cti.com/ca/demos.htm>

<sup>13</sup> En [http://liceu.uab.es/~joaquim/research/Language\\_Resources/SAMPA\\_Catalan.html](http://liceu.uab.es/~joaquim/research/Language_Resources/SAMPA_Catalan.html) se propone una adaptación de SAMPA al catalán.

una fácil edición de las reglas (Pachès *et al.* 2000). La versión actual de SEGRE es el resultado, en el marco del CREL, del esfuerzo conjunto de la UPC (Febrer 2000), la UAB (Riera / de la Mota 2000) y la *Universitat de Barcelona* (UB) (Perea 2000).

También en el contexto del CREL se dispone de un corpus de unidades fonéticas para la síntesis, en el cual se han segmentado y etiquetado casi 800 difonemas, preparado por la UPC y la UAB (Bonafonte / Febrer 2000), y de un corpus prosódico parcialmente etiquetado orientado al modelado de la duración segmental, de las pausas y de los movimientos tonales en la conversión de texto a habla, realizado en colaboración entre la UAB (Riera / Jiménez, 2000), la *Universitat de Girona* (Palmada 2000) y la UPC.

En la antigua Escuela de Ingeniería de Telecomunicación La Salle, actualmente *Enginyeria La Salle* en la URL, surgieron los primeros trabajos sobre conversión de texto a habla en catalán que, a finales de los años 80, dieron lugar a SinCat (Martí 1987) y a Ciber232, comercializado tanto en catalán como en castellano por la empresa Ciberveu, primordialmente para uso de invidentes (Llisterri *et al.* 1993). En esta misma universidad se llevan a cabo también en la actualidad trabajos en el ámbito de la conversión de texto a habla en catalán (Guaus / Iriando 2000 a, b)<sup>14</sup>.

La empresa Telefónica I+D dispone de un conversor multilingüe en castellano, catalán, gallego y vasco (Rodríguez *et al.* 1998), cuya versión catalana se desarrolló a partir de 1997 en colaboración con la UAB. Este sistema se orienta especialmente a los servicios telefónicos.

En el campo del reconocimiento del habla, tanto los grupos como las empresas anteriormente mencionadas cuentan con sistemas que permiten el desarrollo de aplicaciones basadas en un vocabulario específico -por ejemplo, hasta 4000 palabras en el caso del reconocedor de Telefónica I+D<sup>15</sup>- para ofrecer servicios de información o de atención al cliente a través del teléfono.

Los sistemas de reconocimiento requieren también la creación de recursos lingüísticos para su entrenamiento. Cabe destacar el corpus catalán *SpeechDat*, realizado por la UPC en colaboración con la UAB en el marco del CREL<sup>16</sup>. Este corpus, diseñado según los estándares europeos actuales, consta de 1.005 locutores, con 40 elocuciones para cada locutor, y contiene tanto el vocabulario necesario para el diseño de aplicaciones como enunciados que permiten aumentar la cobertura fonética de los programas de reconocimiento (Hernando / Esquerra 2000; Esquerra / Hernando 2000; Mariño *et al.* 2000). Por su parte, Telefónica I+D ha desarrollado, conjuntamente con la UPC, Vocatel, una base de datos en

---

<sup>14</sup> Puede escucharse una demostración en <http://www.salleURL.edu/~iriondo/>

<sup>15</sup> *Reconocedor universal multilingüe*. Telefónica Investigación y Desarrollo. [http://www.tid.es/trabajo/nuestrosproductos/pdfs/servtelefonicos/IV/habla/rec\\_mult.pdf](http://www.tid.es/trabajo/nuestrosproductos/pdfs/servtelefonicos/IV/habla/rec_mult.pdf)

<sup>16</sup> Una descripción detallada se presenta en [http://gps-tsc.upc.es/veu/LR/LR\\_SPD\\_CAT.php3](http://gps-tsc.upc.es/veu/LR/LR_SPD_CAT.php3)

catalán -existente también en gallego- para el reconocimiento de habla en aplicaciones telefónicas (Villarrubia *et al.* 1998; Ezquerro *et al.*, 1998).

El dictado automático -es decir, la transformación del habla a texto escrito- es quizás una de las aplicaciones que mayor interés despierta en el campo del reconocimiento. Con el apoyo de la *Generalitat de Catalunya*<sup>17</sup> y la colaboración de la UPC para el desarrollo técnico y de la UAB y la UB en los aspectos lingüísticos, la empresa Philips lanzó al mercado a finales de 1999 la versión catalana de FreeSpeech 2000, un sistema multilingüe de dictado automático que incorpora, además, vocabularios específicos en los ámbitos legal, médico y de los seguros<sup>18</sup>.

La síntesis y el reconocimiento del habla se integran en los sistemas de diálogo o sistemas conversacionales, que permiten avanzar un paso más en la automatización de determinados servicios, ofreciendo la posibilidad de una interacción entre el usuario y el sistema. Este es tal vez el objetivo al que más esfuerzos se está dedicando hoy en día en el ámbito de las tecnologías del habla y por ello, tanto en las iniciativas institucionales como en las líneas de investigación de los grupos anteriormente mencionados, se han planteado proyectos para el catalán este campo.

Citaremos como ejemplo el prototipo desarrollado conjuntamente por la UPC y la UAB en el contexto del CREL, diseñado para ofrecer un servicio de información telefónica automática a los viajeros de los *Ferrocarrils de la Generalitat* (Nadeu *et al.* 2000; Machuca *et al.* 2000a). El primer paso ha consistido en el análisis de un corpus real de 108 llamadas al servicio de información de los *Ferrocarrils* y en la recogida de un corpus de 200 llamadas a un sistema de diálogo simulado -técnica conocida como "Mago de Oz"- (Machuca *et al.* 2000b), con el fin de conocer las estrategias más adecuadas para la puesta en práctica del sistema, que próximamente empezará a evaluarse. Entre otras iniciativas recientes destaca el proyecto ACIMET, promovido por la *Secretaria per a la Societat de la Informació* del gobierno catalán, cuyo objetivo es ofrecer un servicio telefónico de información meteorológica para el *Servei de Meteorologia* de Cataluña, en el que participan el centro TALP de la UPC, la UAB y la empresa Mensatec Software.

Con esta panorámica que, como se indicaba al principio, no es ni mucho menos completa, hemos pretendido, haciendo hincapié en los proyectos surgidos de la colaboración entre diferentes grupos de I+D, mostrar el desarrollo de las tecnologías del habla en lengua catalana. Al igual que en otras lenguas, los futu-

---

<sup>17</sup> Según los datos recogidos en el capítulo XII del *Informe sobre política lingüística 2000* elaborado por la *Direcció General de Política Lingüística* de la *Generalitat de Catalunya*, la aportación del gobierno catalán fue de 98 millones de pesetas (<http://cultura.gencat.es/llengcat/informe/>).

<sup>18</sup> <http://www.gencat.es/csi/cat/noticies/arxiu/1999/setembre/freespeech.htm>

ros trabajos se encaminarán, muy probablemente, a la mejora de los sistemas de síntesis y de reconocimiento y a la creación de sistemas de diálogo que favorezcan la creación de nuevos servicios -especialmente telefónicos y a través de la web- integrados en una Sociedad de la Información y del Conocimiento cada vez más accesible en la lengua propia de cada ciudadano.

### **Bibliografía**

- Bonafonte, A. / Esquerra, I. / Febrer, A. / Vallverdú, F. (1997): "A Bilingual Text-to-Speech System in Spanish and Catalan", in *Eurospeech'97. 5th European Conference on Speech Communication and Technology*. Rhodes, 22-25 September 1997. Vol. 5. 2455 - 2458.
- Bonafonte, A. / Esquerra, I. / Febrer, A. / Fonollosa, J.A. / Vallverdú, F. (1999): "The UPC Text-to-Speech System for Spanish and Catalan", in *Proceedings of the 5th International Conference on Spoken Language Processing, ICSLP'98*. Sydney, Australia, 30th November-4th December 1998. Vol.5. 1667-1670. <http://gps-tsc.upc.es/veu/research/pubs/download/Bon98c.pdf>
- Bonafonte, A. / Febrer, A. (2000): *Eines de conversió text-parla*, I Jornades del Centre de Referència en Enginyeria Lingüística (CREL), Institut d'Estudis Catalans, Barcelona, 4 i 5 d'abril de 2000. <http://www.crel.net/Jornades/upc/tts.zip>
- Esquerra, I. / Hernando, J. (2000): *SpeechDat: disseny del corpus en català*, I Jornades del Centre de Referència en Enginyeria Lingüística (CREL), Institut d'Estudis Catalans, Barcelona, 4 i 5 d'abril de 2000. <http://www.crel.net/Jornades/upc/speechdat.zip>
- Esquerra, I. / Nadeu, C. / Villarrubia, L. / León, P. (1998): "Design of a phonetic corpus for speech recognition in Catalan", in *Workshop on Language Resources for Minority Languages (LREC'98)*. 27 May 1998, Granada. 89-92.
- Febrer, A. (2000): *Segre: eina de transcripció fonètica*, I Jornades del Centre de Referència en Enginyeria Lingüística (CREL), Institut d'Estudis Catalans, Barcelona, 4 i 5 d'abril de 2000. <http://www.crel.net/Jornades/upc/segre.zip>
- Gaus, R. / Iriondo, I. (2000): "Diphone based Unit Selection for Catalan Text-to-Speech Synthesis", in *TSD 2000, Workshop on Text, Speech and Dialogue*. Brno, September 2000. <http://www.salleURL.edu/~iriondo/publicacions/tsd2000.ps>
- Gaus, R. / Iriondo, I. (2000): "Unit Selection based on Diphones for Catalan Text-to-Speech Conversion", in *Proceedings of the Workshop on Developing Language Resources for Minority Languages: Reusability and Strategic Priorities (LREC-2000)*. Athens, 30 May 2000.
- Hernando, J. / Esquerra, I. (2000): *SpeechDat Catalan Database for Fixed Telephone Network*, I Jornades del Centre de Referència en Enginyeria Lin-

- güística (CREL), Institut d'Estudis Catalans, Barcelona, 4 i 5 d'abril de 2000. <http://www.crel.net/Jornades/upc/speechdat.zip>
- Llisterri, J. / Fernández, N. / Gudayol, F. / Poyatos, J.J. / Martí, J. (1993): "Testing user's acceptance of Ciber232, a text to speech system used by blind persons", in *Speech and Language Technology for Disabled Persons. Proceedings of an ESCA Workshop*. Stockholm, May 31-June 2, 1993. 203-206. [http://liceu.uab.es/~joaquim/publicacions/Stockholm\\_93/stockholm\\_93.html](http://liceu.uab.es/~joaquim/publicacions/Stockholm_93/stockholm_93.html)
- Machuca, M.J. / Bueno, L. / Calonge, R. / Estruch, M. / Riera, M. (2000a): *Eines de reconeixement i prototip de conversa oral*, I Jornades del Centre de Referència en Enginyeria Lingüística (CREL), Institut d'Estudis Catalans, Barcelona, 4 i 5 d'abril de 2000. [http://liceu.uab.es/~joaquim/teaching/speech\\_technology/tecnol\\_parla/dialogue/CREL\\_dialeg/SFI\\_UAB\\_Disseny\\_prototip\\_.htm](http://liceu.uab.es/~joaquim/teaching/speech_technology/tecnol_parla/dialogue/CREL_dialeg/SFI_UAB_Disseny_prototip_.htm)
- Machuca, M.J. / Bueno, L. / Calonge, R. / Estruch, M. / Riera, M. (2000b): *Corpus de diàleg*, I Jornades del Centre de Referència en Enginyeria Lingüística (CREL), Institut d'Estudis Catalans, Barcelona, 4 i 5 d'abril de 2000. [http://liceu.uab.es/~joaquim/teaching/speech\\_technology/tecnol\\_parla/dialogue/CREL\\_dialeg/SFI\\_UAB\\_Corpus\\_Dialeg.htm](http://liceu.uab.es/~joaquim/teaching/speech_technology/tecnol_parla/dialogue/CREL_dialeg/SFI_UAB_Corpus_Dialeg.htm)
- Mariño, J.B. / Padrell, J. / Moreno, A. / Nadeu, C. (2000): "Monolingual and bilingual Spanish-Catalan speech recognizers developed from SpeechDat databases", in *Proceedings of the International Workshop on Very Large Telephone Speech Databases (LREC-2000)*. Athens, May 2000. 57-61. <http://gps-tsc.upc.es/veu/research/pubs/download/Mar00a.pdf>
- Martí, J. (1987): "Un conversor text-veu en català: Sistema SINCAT (SINtetitzador de CATalà)", in Martín Vide, C. (ed.): *Lenguajes naturales y lenguajes formales II*. Barcelona: PPU. 197-209.
- Nadeu, C. / Padrell, J. / Nogueiras, A. / Mariño, J.B. / Bonafonte, A. / Hernandez, J. / Rodríguez, A. / Esquerra, I. / Sesma, A. / Oliver, J. (2000): *Eines de reconeixement i prototip de conversa oral*, I Jornades del Centre de Referència en Enginyeria Lingüística (CREL), Institut d'Estudis Catalans, Barcelona, 4 i 5 d'abril de 2000. <http://www.crel.net/Jornades/upc/prototip-conv.zip>
- Pachès, P. / De La Mota, C. / Riera, M. / Perea, P. / Febrer, A. / Estruch, M. / Garrido, J.M. / Machuca, M.J. / Ríos, A. / Llisterri, J. / Esquerra, I. / Hernandez, J. / Padrell, J. / Nadeu, C. (2000): "SEGRE: An Automatic Tool for Grapheme-to-Allophone Transcription in Catalan", in *Proceedings of the Workshop on Developing Language Resources for Minority Languages: Reusability and Strategic Priorities (LREC-2000)*. Athens, 30 May 2000. 52-61. [http://liceu.uab.es/~joaquim/publicacions/Paches\\_et\\_al\\_2000.pdf](http://liceu.uab.es/~joaquim/publicacions/Paches_et_al_2000.pdf)

- Palmada, B. (2000): *Corpus prosòdic. Etiquetatge ToBI*, I Jornades del Centre de Referència en Enginyeria Lingüística (CREL), Institut d'Estudis Catalans, Barcelona, 4 i 5 d'abril de 2000. [http://www.crel.net/Jornades/UdG/ToBI\\_CREL.zip](http://www.crel.net/Jornades/UdG/ToBI_CREL.zip)
- Perea, P. (2000): *SEGRE: una eina de transcripció aplicada a las varietats dialectals del català*, I Jornades del Centre de Referència en Enginyeria Lingüística (CREL), Institut d'Estudis Catalans, Barcelona, 4 i 5 d'abril de 2000. <http://www.crel.net/Jornades/UB/segre-p.zip>
- Riera, M. / De La Mota, C. (2000): *El transcriptor automàtic SEGRE*, I Jornades del Centre de Referència en Enginyeria Lingüística (CREL), Institut d'Estudis Catalans, Barcelona, 4 i 5 d'abril de 2000. [http://liceu.uab.es/~joaquim/teaching/speech\\_technology/tecnol\\_parla/synthesis/transcription/CREL/SFI\\_UAB\\_Transcriptor.pdf](http://liceu.uab.es/~joaquim/teaching/speech_technology/tecnol_parla/synthesis/transcription/CREL/SFI_UAB_Transcriptor.pdf)
- Riera, M. / Jiménez, E. (2000): *Corpus prosòdic*, I Jornades del Centre de Referència en Enginyeria Lingüística (CREL), Institut d'Estudis Catalans, Barcelona, 4 i 5 d'abril de 2000. [http://liceu.uab.es/~joaquim/teaching/Language\\_resources/spoken\\_res/CREL/SFI\\_UAB\\_Corpus\\_Prosodic.pdf](http://liceu.uab.es/~joaquim/teaching/Language_resources/spoken_res/CREL/SFI_UAB_Corpus_Prosodic.pdf)
- Rodríguez, M.A. / Escalada, J.G. / Torre, D. (1998): "Conversor texto-voz multilingüe para español, catalán, gallego y euskera", *Procesamiento del Lenguaje Natural* 23: 16-23.
- Villarrubia, L. / León, P. / Hernández, L. / Nadeu, C. / Esquerra, I. / Hernando, J. / García Mateo, C. / Docío, L. (1998): "VOCATEL and VOGATEL: Two Telephone Speech Databases of Spanish Minority Languages (Catalan and Galician)", in *Proceedings of the Workshop on Language Resources for European Minority Languages (LREC -98)*. Granada, 27 May 1998. 1-4.