



Las tecnologías del habla

JOAQUIM LLISTERRI

Departament de Filologia Espanyola
Universitat Autònoma de Barcelona

Joaquim.Llisterri@uab.cat
<http://liceu.uab.cat/~joaquim>

Materiales en la web:

http://liceu.uab.cat/~joaquim/speech_technology/UAM_07/UAM_07.html

GUIÓN

1. LAS TECNOLOGÍAS DEL HABLA

La interacción oral con los sistemas informáticos: nuevas necesidades y nuevas tecnologías. Salida vocal: conversión de texto en habla (TTS, *Text-to-Speech Synthesis*). Entrada vocal: reconocimiento automático del habla (ASR, *Automatic Speech Recognition*). Interacción con el usuario: sistemas de diálogo o sistemas conversacionales (SLS, *Spoken Language Systems*).

2. LA CONVERSIÓN DE TEXTO EN HABLA

Esquema general de un sistema de conversión de texto en habla. Etapas en la conversión de texto en habla: procesamiento previo del texto, transcripción fonética automática, análisis lingüístico, asignación de elementos prosódicos –duración, intensidad, pausas y entonación-, selección de unidades y conversión en valores de parámetros de síntesis. Aplicaciones de la conversión de texto en habla.

3. EL RECONOCIMIENTO AUTOMÁTICO DEL HABLA

Esquema general de un sistema de reconocimiento automático del habla. Etapas en el reconocimiento del habla: entrenamiento, parametrización de la señal y comparación con los modelos almacenados. Aplicaciones del reconocimiento automático del habla: reconocimiento de palabras aisladas y reconocimiento de habla continua. Identificación y la verificación automáticas del locutor.

4. LOS SISTEMAS DE DIÁLOGO

Módulos de un sistema de diálogo: reconocimiento automático del habla, comprensión del habla natural (SLU, *Spoken Language Understanding*), gestión del diálogo, bases de datos, generación de lenguaje natural (NLG, *Natural Language Generation*), conversión de texto en habla. Aplicaciones de los sistemas de diálogo: sistemas de información y sistemas de transacciones. Sistemas multimodales. Traducción automática del habla.



BIBLIOGRAFÍA BÁSICA

TECNOLOGÍAS DEL HABLA

COLEMAN, J. (2005) *Introducing Speech and Language Processing*. Cambridge: Cambridge University Press (Cambridge Introduction to Language and Linguistics) (+CD-ROM). <http://www.islp.org.uk/>

CORTÁZAR, I.- RODRÍGUEZ, M.A.- GARRIDO, J.M.- CAMINERO, F.J.- BERNAT, J.- RELAÑO, J.- GARIJO, F.J.- HERNÁNDEZ, L.A. (2002) "Últimos desarrollos en tecnologías de voz y del lenguaje", *Comunicaciones de Telefónica I+D* 24: 25-64. http://www.tid.es/documentos/revista_comunicaciones_i%2Bd/numero_24.pdf

HOLMES, J.N.- HOLMES, W. (2001) *Speech Synthesis and Recognition*. London: Taylor & Francis, 2nd edition.

JURAFSKY, D.- MARTIN, J.H. (2000) *Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition*. New Jersey: Prentice Hall. <http://www.cs.colorado.edu/~martin/slp.html>

KELLER, E. (Ed.) (1994) *Fundamentals of Speech Synthesis and Speech Recognition. Basic Concepts, State of the Art and Future Challenges*. Chichester: John Wiley & Sons.

LLISTERRI, J. (2001) "Les technologies de la parla", in MARTÍ, M.A. (Coord.) *Les tecnologies del llenguatge*. Barcelona: Edicions de la Universitat Oberta de Catalunya (Manuals, 53). pp. 239-272; trad. cast.: "Las tecnologías del habla", in MARTÍ, M.A. (Coord.) *Tecnologías del lenguaje*. Barcelona: Editorial UOC (Manuales, Humanidades), 2003. pp. 249-281.

LLISTERRI, J.- CARBÓ, C.- MACHUCA, M. J.- de la MOTA, C.- RIERA, M.- RÍOS, A. (2003) "El papel de la lingüística en el desarrollo de las tecnologías del habla", in CASAS GÓMEZ, M. (Dir.) - VARO VARO, C. (Ed.) *VII Jornadas de Lingüística*. Cádiz: Servicio de Publicaciones de la Universidad de Cádiz. pp. 137-191. http://liceu.uab.es/~joaquim/publicacions/Llisterri_Carbo_Machuca_Mota_Riera_Rios_03_Linguistica_Tecnologias_Habla.pdf

O'SHAUGHNESSY, D. (1987) *Speech Communication. Human and Machine*. Reading, Mass.: Addison Wesley. Second Edition: IEEE Press, 2000.

PARDO, J.M. (2004) "Tecnología del Habla: Debate sobre retos pendientes y perspectivas", *Procesamiento del Lenguaje Natural* 32: 37-42. <http://www-gth.die.upm.es/research/documentation/AN-44Tec-04.doc>

SCHROEDER, M.R. (2004) *Computer Speech: Recognition, Compression, Synthesis*. Second Edition. Berlin: Springer-Verlag.

SCHULTZ, T.- KIRCHHOFF, K. (Eds.) (2006) *Multilingual Speech Processing*. San Diego - London: Academic Press.

VILLARRUBIA, L.- GARRIDO, J.M.- RELAÑO, J.- CAMINERO, J.- ESCALADA, J.G.- RODRÍGUEZ, M.C.- HERNÁNDEZ, L.A. (2002) "Productos de tecnología del habla para Latinoamérica", *Comunicaciones de Telefónica I+D* 27: 53-72. http://www.tid.es/documentos/revista_comunicaciones_i%2Bd/numero_27.pdf

VILLARRUBIA, L.- RODRÍGUEZ, A.- RELAÑO, J.- GARIJO, F.J.- BERNAT, J.- HERNÁNDEZ, L.A.- TAPIAS, D.- MARÍA, L.A. (2003) "Tecnología del habla para



aplicaciones multilingües, multiservicio y multiplataforma", *Comunicaciones de Telefónica I+D* 30: 47-78.

http://www.tid.es/documentos/revista_comunicaciones_i%2Bd/numero_30.pdf

ZUE, V. (2004) "Eighty challenges facing speech input/output technologies", in *From Sound to Sense: 50+ Years of Discoveries in Speech Communication*. Massachusetts Institute of Technology, Cambridge, MA., 11-13 June 2004.

<http://www.rle.mit.edu/soundtosense/conference/pdfs/invitedspeakers/Zue%20PAPER.pdf>

CONVERSIÓN DE TEXTO EN HABLA

BONAFONTE, A.- ESCUDERO, D.- RIERA, M. (2006) "La conversión de texto en habla", in LLISTERRI, J.- MACHUCA, M. J. (Eds.) *Los sistemas de diálogo*. Bellaterra - Soria: Universitat Autònoma de Barcelona, Servei de Publicacions - Fundación Duques de Soria (Manuales de la Universitat Autònoma de Barcelona, Lingüística, 45). pp. 177-208.

DUTOIT, T. (1997) *An Introduction to Text-to-Speech Synthesis*. Dordrecht: Kluwer Academic Publishers (Text, Speech and Language Technology, 3).

<http://tcts.fpms.ac.be/publications/books/introtts/>

LLISTERRI, J.- CARBÓ, C.- MACHUCA, M. J.- de la MOTA, C.- RIERA, M.- RÍOS, A. (2004) "La conversión de texto en habla: aspectos lingüísticos", in MARTÍ, M. A. – LLISTERRI, J. (Eds.) *Tecnologías del texto y del habla*. Barcelona. Edicions de la Universitat de Barcelona – Fundación Duques de Soria (UB, 72). pp. 145-186.

http://liceu.uab.cat/~joaquim/publicacions/Llisterri_Carbo_Machuca_Mota_Riera_Rios_04_Conversion_Texto_Habla.pdf

NUSBAUM, H. C. - SHINTEL, H. (2006) "Speech synthesis", in BROWN, K. (Ed.) *Encyclopedia of Language & Linguistics*. Amsterdam: Elsevier. pp. 19-31.

<http://dx.doi.org/10.1016/B0-08-044854-2/00913-5>

SCHROETER, J. (2006) "Text-to-Speech (TTS) Synthesis", in DORF, R.C. (Ed.) *The Electrical Engineering Handbook*. 3rd edition. Boca Raton, FL: CRC Press. pp. 16-1 - 16-13.

http://www.research.att.com/~ttsweb/tts/papers/2005_EEHandbook/tts.pdf

RECONOCIMIENTO AUTOMÁTICO DEL HABLA

DEROO, O. (1999) *A Short Introduction to Speech Recognition*. TCTS Lab, Faculté Polytechnique de Mons.

<http://tcts.fpms.ac.be/asr/intro.php>

HATON, J.P.- CERISARA, C.- FOHR, D.- LAPRIE, Y.- SMAÏLI, K. (2006) *Reconnaissance automatique de la parole. Du signal à son interprétation*. Paris: Dunod (UniverSciences).

<http://parole.loria.fr/livreParole/>

JUANG, B.-H. - RABINER, L. R. (2006) "Automatic speech recognition: History", in BROWN, K. (Ed.) *Encyclopedia of Language & Linguistics*. Amsterdam: Elsevier. pp. 806-819.

<http://dx.doi.org/10.1016/B0-08-044854-2/00906-8>

Speech Recognition Systems. CALL Information Sheet 15. Revised January 2005. Edinburgh: CALL Centre, University of Edinburgh.

http://callcentre.education.ed.ac.uk/downloads/speech_recognition/introduction.pdf



TORRES, M. I. (2006) "El reconocimiento del habla", in LLISTERRI, J.- MACHUCA, M. J. (Eds.) *Los sistemas de diálogo*. Bellaterra - Soria: Universitat Autònoma de Barcelona, Servei de Publicacions - Fundació Duques de Soria (Manuals de la Universitat Autònoma de Barcelona, Lingüística, 45). pp. 81-98.

SISTEMAS DE DIÁLOGO

LLISTERRI, J. (2006) "Introducción a los sistemas de diálogo", in LLISTERRI, J.- MACHUCA, M. J. (Eds.) *Los sistemas de diálogo*. Bellaterra - Soria: Universitat Autònoma de Barcelona, Servei de Publicacions - Fundació Duques de Soria (Manuals de la Universitat Autònoma de Barcelona, Lingüística, 45). pp. 11-21.

http://liceu.uab.cat/~joaquim/publicacions/Llisterri_06_Sistemas_Dialogo.pdf

LLISTERRI, J.- MACHUCA, M. J. (Eds.) (2006) *Los sistemas de diálogo*. Bellaterra - Soria: Universitat Autònoma de Barcelona, Servei de Publicacions - Fundació Duques de Soria (Manuals de la Universitat Autònoma de Barcelona, Lingüística, 45).

http://liceu.uab.cat/~joaquim/publicacions/Llisterri_Machuca_06_Sistemas_Dialogo.pdf

LÓPEZ-CÓZAR, R. (2006) "Análisis y gestión del diálogo", in LLISTERRI, J.- MACHUCA, M. J. (Eds.) *Los sistemas de diálogo*. Bellaterra - Soria: Universitat Autònoma de Barcelona, Servei de Publicacions - Fundació Duques de Soria (Manuals de la Universitat Autònoma de Barcelona, Lingüística, 45). pp. 119-152.

LÓPEZ-CÓZAR, R. - ARAKI, M. (2005) *Spoken, Multilingual and Multimodal Dialogue Systems: Development and Assessment*. Chichester: John Wiley & Sons.

MINKER, W.- BENNACEF, S. (2004) *Speech and Human-Machine Dialog*. Dordrecht: Kluwer (The Kluwer International Series in Engineering and Computer Science, 770).

TAPIAS, D. (2002) "Interfaces de voz con lenguaje natural", in MARTÍ, M. A. - LLISTERRI, J. (Eds.) (2002) *Tratamiento del lenguaje natural. Tecnología de la lengua oral y escrita*. Barcelona: Edicions Universitat de Barcelona - Fundació Duques de Soria (Biblioteca de la Universitat de Barcelona, Manuales, 53). pp. 189-207.

WAIBEL, A. (2001) "Los sistemas integrales completos del habla, del lenguaje y la interfaz humana", *Quark. Ciencia, Medicina, Comunicación y Cultura* 21: 95-102.

<http://www.prbb.org/quark/21/021095.htm>