



La ingeniería lingüística en la sociedad de la información

Maria Antònia Martí Antonín



Profesora del Departamento de Lingüística General y directora del Centro de Lenguaje y Computación (CLiC) de la Universidad de Barcelona
amarti@lingua.fil.ub.es

Joaquim Llisterri



Profesor del Departamento de Filología Española de la Universidad Autónoma de Barcelona
Joaquim.Llisterri@uab.es

Resumen: Las necesidades que plantea la sociedad de la información determinan en cierta medida los objetivos de la lingüística computacional en el cambio de siglo: la revolución tecnológica ha dado lugar a cambios en profundidad en los sectores productivos y las profesiones relacionadas con el lenguaje, y también en aquellos sectores que, sin estar directamente relacionados con el lenguaje, necesitan dotar su actividad del valor añadido de una comunicación eficiente y fluida con clientes y usuarios. Esta nueva realidad impone la necesidad de desarrollar aplicaciones de procesamiento del lenguaje y del habla que tendrán un amplio abanico de posibles usuarios y que han de tratar con centenares de millones de documentos escritos o sonoros, lo que a su vez implica la necesidad de disponer de recursos básicos de ingeniería lingüística que deben permitir el desarrollo efectivo de estas aplicaciones. En este artículo se presentan las principales tecnologías lingüísticas, las relativas al tratamiento tanto del habla como del texto escrito, y su papel en el contexto europeo actual.

1. Comunicación y lenguaje en el marco de la sociedad de la información

El lenguaje es el medio principal a través del cual los humanos intercambiamos información. Ningún otro sistema de información es tan eficaz para compartir ideas, transmitir conocimientos, expresar emociones y crear, en definitiva, lo que conocemos como cultura. Así, no es extraño que se relacione la aparición del lenguaje con la formación de grupos sociales cohesionados y con la necesidad de disponer de un procedimiento que permita la expresión del conocimiento común. *Lenguaje, sociedad e información* son tres términos que aparecen estrechamente ligados desde el inicio de la civilización.

En su sentido más actual, el concepto de *sociedad de la información* incorpora otros componentes, entre los cuales hay que destacar las tecnologías que permiten la difusión del conocimiento compartido. No es la primera vez que nos encontramos ante técnicas que facilitan la transmisión del saber; sólo hay que recordar que el papel de las nuevas tecnologías en la pervivencia de las lenguas y las culturas a menudo ha sido comparado con el que en su momento tuvo la imprenta (Danzin, 1992).



Las nuevas tecnologías ponen a nuestro alcance una gran variedad de datos y de modos de interacción: en Internet podemos acceder a sistemas de enseñanza asistidos por ordenador, comprar, establecer nuevas relaciones, conversar con amigos que están al otro lado del planeta y acceder a todo tipo de información. Puede decirse que el acceso a la información se ha "democratizado": antes de Internet, si queríamos enterarnos de la investigación que se desarrollaba en una universidad, era necesario ir al centro y permanecer en él un tiempo con el objeto de conocer las líneas de investigación y a los investigadores implicados. Ahora, podemos acceder a cualquier campus universitario, obtener esta información e incluso imprimir sus publicaciones mediante una simple consulta a Internet, sin movernos del lugar donde trabajamos.

Podemos afirmar que se está creando un nuevo nivel de realidad, la realidad virtual, que nos permite realizar interacciones comunicativas equivalentes a las que tienen lugar en la realidad presencial, aunque desligadas de muchos de los condicionamientos que nos imponen los parámetros espacio-tiempo. Ahora bien, el nuevo entorno comunicativo que deriva de la aplicación de estas nuevas tecnologías plantea nuevos problemas: el acceso a la información debe ser eficaz, rápido y sencillo, y hay que prever la posibilidad de que los usuarios cometan errores. También es deseable que en el mismo entorno se integre la voz, la imagen y la lengua escrita.

La aparición de Internet como red de telecomunicaciones y de la WWW cambia el concepto de lo que se considera información: no está mejor informado quien tiene más datos, sino quien dispone de los mejores medios para obtener exclusivamente los que necesita.

La recuperación de información, un problema que hasta hace poco afectaba a colectivos de profesionales muy específicos, ha pasado a ser uno de los problemas clave que deberá afrontar la sociedad de la información. Internet constituye un gran banco de datos que puede volverse inservible si no se dispone de sistemas de recuperación de información que satisfagan las demandas de los usuarios. Además, las instituciones, los hospitales, las entidades bancarias y, en general, las empresas que generan una gran cantidad y variedad de datos también necesitan sistemas que les permitan gestionarlos de forma eficaz.

En consecuencia, es preciso diseñar sistemas informáticos que, ante una demanda del usuario, recuperen los documentos que responden a sus objetivos de información y que ignoren los irrelevantes. Además, hay que hacer frente al problema del multilingüismo, tanto en las demandas del usuario como en los textos que se recuperan.

Las necesidades que plantea la sociedad de la información determinan en cierta medida los objetivos de la lingüística computacional en el cambio de siglo: la necesidad de tratar las producciones lingüísticas sin restricciones constituye el aspecto clave del cambio que se produce actualmente. Las repercusiones de este nuevo enfoque afectan en profundidad la definición de los campos de investigación tradicionales. Por primera vez hay que desarrollar aplicaciones de procesamiento del lenguaje y del habla que tendrán un amplio abanico de posibles usuarios y que tienen que tratar con centenares de millones de documentos escritos o sonoros. Esta nueva realidad impone la necesidad de disponer de recursos de ingeniería lingüística básicos que deben permitir el desarrollo efectivo de



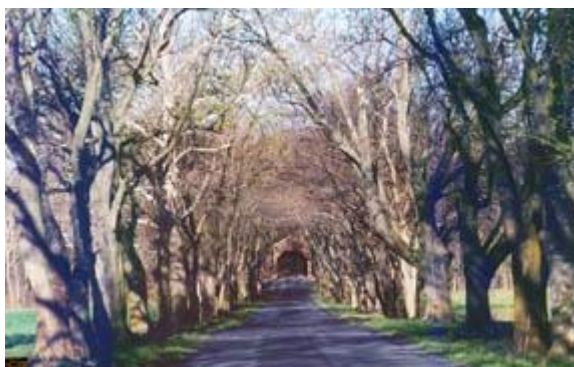
estas aplicaciones.

Finalmente, la revolución tecnológica ha dado lugar a cambios en profundidad en los sectores productivos y las profesiones relacionadas con el lenguaje. El mundo de la edición y la traducción se enfrenta a la necesidad de definir nuevos perfiles profesionales y de renovar las estrategias productivas y comerciales.

2. Las tecnologías lingüísticas

En el marco que acabamos de describir, las interacciones comunicativas no pueden concebirse sin el complemento de unas tecnologías, llamadas "de la información y la comunicación", que se apoyan, en buena medida, en el uso de sistemas informáticos y de datos lingüísticos debidamente formalizados.

El procesamiento de los textos y del habla constituye el núcleo de las actuales tecnologías del lenguaje, donde confluyen la informática y la lingüística, hecho que explica que en el ámbito de la filología se utilice también el término *lingüística informática*. Una de las acciones del programa europeo Tecnologías para la Sociedad de la Información ([IST, Information Society Technologies](#)) en el [V Programa marco](#) reconoce explícitamente el papel de las tecnologías lingüísticas en el desarrollo de la sociedad de la información, que incluye, entre sus acciones clave, la línea denominada Tecnologías para el lenguaje humano ([HLT, Human Language Technologies](#)).



Estas tecnologías constituyen también la base de lo que se llama *ingeniería lingüística*, entendida como la aplicación de los conocimientos sobre la lengua en el desarrollo de sistemas informáticos que puedan reconocer, comprender, interpretar y generar lenguaje humano en todas sus formas (véase [Ingeniería lingüística. Cómo aprovechar la fuerza del lenguaje](#)).

Las tecnologías lingüísticas engloban una serie de técnicas relacionadas con el tratamiento informático del lenguaje, y se suelen distinguir, tal como se presenta a continuación, las que se aplican al tratamiento de la lengua hablada de las que son propias del tratamiento del texto escrito.

3. Las tecnologías del texto escrito

En el momento actual, el acceso a la información debe ser selectivo y guiado por objetivos; si no es así, a pesar de disponer de una gran cantidad de datos a nuestro alcance, no se llegarán a convertir en información que podamos integrar como conocimiento. Este hecho plantea la necesidad de definir métodos de almacenaje y codificación de la información a fin de que posteriormente pueda ser debidamente recuperada.

Actualmente, en estos entornos los documentos se indizan por formas o por palabras clave, sin tener



en cuenta las relaciones entre las variantes flexivas de las palabras o las relaciones de significado. Tanto las empresas editoriales, en cuanto creadoras de contenidos estructurados (diccionarios y enciclopedias), como los sistemas de búsqueda en Internet o la gestión de los grandes bancos de datos requieren sistemas que permitan el acceso al contenido desde diferentes perspectivas según las necesidades de los usuarios.



El desarrollo de estos sistemas requiere disponer de recursos de ingeniería lingüística que permitan optimizar los resultados de las búsquedas. Las mejoras se pueden producir en diferentes sentidos:

- Permitiendo expresar las búsquedas en lenguaje natural y sin restringirlas a un metalenguaje, normalmente una combinación de palabras clave y operadores booleanos, que es necesario aprender.
- Ampliando el espacio de búsqueda mediante la ampliación de la pregunta del usuario incorporando en él sus formas flexivas y las palabras semánticamente relacionadas.
- Restringiendo los documentos seleccionados a aquellos que realmente responden al sentido de las búsquedas.

Entre los [recursos de ingeniería lingüística](#), destacan:

- Los analizadores morfológicos, que permiten relacionar las variantes flexivas de una palabra con su lema y asignan una descripción morfológica a las palabras de los textos.
- Los desambiguadores, que, ante una ambigüedad en el nivel morfológico, sintáctico o semántico, deciden, a partir de datos estadísticos o sobre la base de reglas lingüísticas, cuál es la interpretación más correcta.
- Los analizadores sintácticos, que identifican los grupos de palabras que funcionan como un todo, lo que permite identificar las entidades del texto.
- Las redes semánticas, que organizan el léxico según sus relaciones internas y permiten etiquetar semánticamente los textos.
- Los corpus etiquetados morfológicamente, sintácticamente y semánticamente, que se utilizan como fuentes de información para el aprendizaje de los etiquetadores automáticos y como consulta para la búsqueda lingüística. La utilización de corpus va ligada a una serie de procesos —como la codificación, el etiquetado y el análisis lingüístico— y herramientas —como por ejemplo los programas de concordancias— que facilitan su uso y explotación.



- Lexicones computacionales que guían el proceso de análisis e interpretación semántica del texto.
- Lexicones multilingües que permiten el proceso de traducción automática o bien ayudan en la traducción manual.

Las técnicas asociadas a la gestión de contenidos utilizan estos recursos como fuentes de información para el proceso de los datos. Las técnicas más significativas son la extracción de información, la recuperación de información, la traducción automática y la elaboración de resúmenes.

El objetivo último es que los documentos puedan ser indizados por conceptos, que se pueda acceder a la información mediante interacciones en lenguaje natural y que podamos interactuar y recuperar la información en la lengua que deseemos.

4. Las tecnologías del habla

El uso de la lengua oral como modo de comunicación con los sistemas informáticos permite, en primer lugar, liberar las manos y la vista, lo que facilita la recepción de la información desde cualquier punto y hace posible la movilidad, de manera que pueden hacerse otras actividades simultáneas. La entrada de datos sobre productos en la gestión de los fondos de un almacén, las operaciones relacionadas con el mantenimiento y el control de la calidad en entornos industriales, la recepción de información en entornos como cabinas de avión —con una gran cantidad de sistemas de gobierno e indicadores—, el control de la radio, del teléfono o de la climatización de un coche y la obtención de información sobre itinerarios —el prototipo desarrollado en el proyecto [Vodis \(Advanced Speech Technologies for Voice-operated Driver Information Systems\)](#) ofrece estas posibilidades—, o bien [el dictado de documentos](#) son algunas de las aplicaciones para las cuales se ha demostrado que utilizar la lengua hablada es una ventaja importante.

En segundo lugar, utilizar el habla para acceder a la información almacenada en un sistema informático nos brinda la posibilidad de usar el teléfono en vez de la pantalla y el teclado tradicionales, con todo lo que ello implica en cuanto a rapidez; hay que pensar, además, en la independencia que este procedimiento proporciona con respecto a la distancia y la disponibilidad de un ordenador. Algunas aplicaciones en este ámbito son los servicios de información sobre horarios de transportes públicos, las consultas en las páginas amarillas u otros directorios, la información ciudadana, la banca telefónica o las centralitas automatizadas.

Finalmente, no hay que olvidar que guardar y recibir información mediante el habla es una alternativa práctica cuando hay que prescindir, debido al tamaño, del teclado, como sucede con las agendas electrónicas —productos como [Artcommander](#) facilitan la entrada de datos por voz— y los comunicadores personales. La voz, junto con los lápices electrónicos y las pantallas táctiles, es una modalidad viable para hacer aumentar la funcionalidad de estos aparatos.





Hoy en día disponemos de programas comerciales que posibilitan el dictado automático —como por ejemplo los desarrollados por [Dragon Systems](#), [IBM](#) o [Philips](#)—, la navegación por Internet —el [Free Speech Browser](#) sirve de ejemplo— y la recuperación de información mediante el habla. También es posible la lectura en voz alta de un texto en formato electrónico, como podría ser un mensaje de correo —por ejemplo, el proyecto [EMU](#)— o el contenido de una base de datos. Es posible, además, con los sistemas de diálogo que incorporan síntesis, reconocimiento y un cierto grado de comprensión, obtener información o efectuar transacciones comerciales mediante el teléfono.

Son seguramente estos últimos sistemas los que en la actualidad despiertan más interés, y por este motivo mencionaremos algunas de sus aplicaciones.

Uno de los sectores para los que se han desarrollado sistemas de diálogo es el de los transportes públicos, ya que a menudo las centralitas telefónicas no pueden atender todas las llamadas de los usuarios, especialmente las referentes a consulta de horarios. Proyectos como [ARISE \(Automatic Railway Information Systems for Europe\)](#) han desarrollado prototipos para dar respuesta a este tipo de peticiones, concretamente en el caso de los ferrocarriles, mientras que, por ejemplo, [ATIS \(Advanced Travel Information System\)](#) y [Pegasus](#) proporcionan información sobre horarios de vuelos.

Existen también sistemas de diálogo complementados con información multimodal, entre los que puede mencionarse [Waxholm](#), que ofrece información sobre el transporte marítimo en el archipiélago de Estocolmo utilizando la voz acompañada de un rostro generado también mediante ordenador; [MASK](#), un quiosco que proporciona datos sobre circulación de trenes franceses tanto verbalmente como mediante una pantalla y que permite también reservar billetes, o [TOOT](#), diseñado para realizar consultas vocales en la web y obtener horarios de trenes.

La información ciudadana es otro servicio que puede ofrecerse de forma automática mediante el teléfono. [Voyager](#), por ejemplo, informa en inglés, francés e italiano sobre direcciones, números de teléfono y localización de restaurantes, hoteles, oficinas de correos, estaciones de metro, etc., en el área de Boston. De esta misma zona se puede recibir información sobre restaurantes con el sistema [DINEX](#): un teléfono y un ordenador conectados a Internet permiten saber la localización, el precio, el tipo de tarjeta de crédito aceptado y, en algunos casos, el menú, todo ello haciendo preguntas en lenguaje natural como por ejemplo "¿Qué restaurantes chinos hay cerca del Hotel Plaza?". Existen también aplicaciones, como [Movieline](#) —centrada en la cartelera de los cines de Pittsburgh—, que ofrecen información hablada.

En otros ámbitos como el de la información meteorológica destacan sistemas como [Jupiter](#), que proporciona de manera automática y mediante el teléfono datos sobre el tiempo de más de quinientas ciudades del mundo.

Por último, debe señalarse la importancia que está adquiriendo la integración entre los sistemas de diálogo y la web, tal como se hace evidente en [DINEX](#). Uno de los proyectos más ambiciosos en este campo es [WebGALAXY](#), orientado a la búsqueda de información en Internet mediante preguntas libres, tratadas por un reconocedor de habla y por un sistema de comprensión del lenguaje natural, complementados por un procedimiento de navegación.

Una aplicación que cada vez tiene una importancia mayor son los sistemas de traducción automática de conversaciones, tanto telefónicas como cara a cara. Proyectos como [Verbmobil](#), [JANUS](#) o [NESPOLE \(Negociating through Spoken Language in E-commerce\)](#), este último especialmente orientado al comercio electrónico, ponen de manifiesto el futuro de un campo que requiere simultáneamente tecnologías propias del procesamiento del lenguaje natural, como la traducción automática, y tecnologías del habla, como el reconocimiento, la síntesis y el diálogo.

La combinación de estas tecnologías es también indispensable en la enseñanza de lenguas asistida por ordenador, tanto en los sistemas presenciales como en la enseñanza a distancia. Entre los primeros pueden mencionarse proyectos como [ISLE \(Interactive Spoken Language Education\)](#), aplicaciones como las desarrolladas en [SRI International](#) o productos como los que comercializa [Auralog](#), orientados a la enseñanza de lenguas extranjeras; para los niños que aprenden la lengua materna, herramientas como [LISTEN](#) facilitan la práctica de la lectura. En la enseñanza a



distancia son útiles, en el campo que nos ocupa, sistemas como [WebGrader](#), que posibilitan evaluar la pronunciación mediante la web.

También hay que referirse al papel de las tecnologías lingüísticas como ayuda para los colectivos con necesidades especiales. Los [conversores de texto a habla](#) o los sistemas que combinan un escáner con un conversor, como el ya clásico [Kurzweil](#), posibilitan la lectura a los invidentes. Las personas con dificultades de audición encuentran en la conversión de habla a texto mediante técnicas de reconocimiento una ayuda importante si disponen de teléfonos que incorporen pantallas, y también pueden utilizar la lectura labial mediante un rostro generado artificialmente, tal como muestra el proyecto [Teleface](#). Las limitaciones en la producción del habla pueden resolverse con sistemas de generación de mensajes y, en el caso de personas con problemas de movilidad, el control del entorno doméstico mediante la voz aumenta notablemente su independencia.

5. La sociedad de la información en el contexto europeo

Podría decirse que el inicio del [V Programa marco de investigación y desarrollo](#) (1998-2002) situó el concepto de sociedad de la información en una posición central para el desarrollo de la Unión Europea.

Tomando como puntos de referencia documentos como el [Green paper on living and working in the information society: people first](#), se llegó, en enero de 1999, a la decisión del Consejo de adoptar un nuevo programa denominado Tecnologías para la sociedad de la información ([IST, Information Technologies Society](#)). El programa se estructura en diversas acciones clave, entre las cuales hay una dedicada [a las herramientas y los contenidos multimedia](#); en esta acción se enmarca una línea centrada en las tecnologías lingüísticas ([HLT, Human Language Technologies](#)), cuyos contenidos describiremos muy brevemente a continuación.

Al mismo tiempo, diversos países —entre ellos [España](#)— han desarrollado programas que tienen como objetivo promover la sociedad de la información, complementando en el ámbito nacional las iniciativas europeas. Cataluña también tiene un plan para hacer más accesible a todo el mundo la sociedad de la información, gestionado por la [Secretaría para la Sociedad de la Información](#) del [Departamento de Universidades, Investigación y Sociedad de la Información](#).

La definición del papel de las tecnologías lingüísticas en la sociedad de la información se inició, en el contexto europeo, en el año 1997 con el documento [Human language technologies. Living and working together in the information society](#), en el que se definen las grandes líneas que configuran el programa actual: multilingüismo, interacción natural y contenidos activos. Como resultado de consultas a diversos sectores, el informe [Human language, human communication, human knowledge](#) precisaba de una forma más detallada los futuros contenidos de esta línea de acción, para llegar, finalmente, a la publicación del programa de trabajo en 1999, actualizado posteriormente para el [año 2000](#) (actualmente está en preparación el plan de trabajo del [año 2001](#)).

Uno de los principales objetivos del programa es ayudar a superar las barreras lingüísticas para poder acceder con más facilidad a los contenidos y a los servicios digitales, como por ejemplo las transacciones a distancia, la gestión distribuida de conocimiento corporativo o los servicios de información en línea. Ello requiere desarrollar herramientas de análisis, traducción, resumen y generación de contenidos lingüísticos, y tecnologías —agentes de búsqueda y de presentación de contenidos— que permitan la recuperación y la presentación de información en diversas lenguas. También se prevé en esta línea de acción la transferencia de tecnologías que parecen especialmente interesantes a una gama más amplia de lenguas.

Desde un punto de vista práctico, el desarrollo de proyectos enmarcados en este ámbito debe permitir la automatización de las centralitas telefónicas y los servicios de atención al cliente, la identificación automática de la lengua y la verificación de la identidad del usuario a partir de su voz, el filtrado, la traducción y la respuesta de correo electrónico, la gestión del acceso a catálogos multilingües o la oferta de servicios comerciales en la web con una interfaz oral. En el terreno de la lengua escrita, la



autoría conjunta de documentos, el desarrollo de herramientas de corrección para hablantes no nativos, las herramientas de gestión de documentos y de terminología técnica en diversas lenguas o los programas de enseñanza de lenguas asistida por ordenador serán algunos desarrollos que el programa hará posibles. Por último, también se plantea la creación de servicios en Internet que tengan en cuenta las preferencias lingüísticas de los usuarios y que permitan crear contenidos publicados simultáneamente en más de una lengua y adaptados a contextos culturales diferentes.

La segunda línea de acción, denominada genéricamente *interactividad natural*, se centra en conseguir que los sistemas informáticos se puedan utilizar más fácilmente mediante el uso espontáneo de la lengua hablada o escrita, de forma que superen las limitaciones actuales en nuestro modo de interacción con los ordenadores. Se trata, en conjunto, de avanzar hacia el diseño de interfaces más intuitivas, multilingües y multimodales que permitan, por ejemplo, la navegación por la web, el control de sistemas o el diálogo persona-máquina, gracias a la comprensión de información hablada, escrita o, incluso, gestual.

Las tecnologías necesarias para hacer posible estas aplicaciones son, esencialmente, el reconocimiento y la síntesis del habla de alta calidad, y la comprensión de la lengua oral en situaciones de interacción —podríamos decir de conversación— entre un usuario y un sistema informático.

Finalmente, una tercera línea de actuación se centra en los contenidos, especialmente en lo que concierne a la gestión, búsqueda y recuperación de información multilingüe. Para alcanzar este objetivo se prevé la utilización de agentes lingüísticos inteligentes y de técnicas que permitan identificar el tema de los textos, extraer datos y clasificarlos automáticamente, con el fin de crear, resumir, convertir y presentar información en diversas lenguas, tanto en su representación escrita como oral, incluyendo diálogos y componentes multimedia.

Puede observarse, por lo tanto, que las líneas de acción europeas enmarcan las tecnologías lingüísticas en el contexto de la sociedad de la información, insistiendo especialmente en los nuevos retos que hemos querido destacar en este trabajo, centrados en el procesamiento multilingüe de la información, tanto escrita como oral.

6. Perspectivas profesionales

La revolución tecnológica que ha significado la incorporación de estos nuevos sistemas de comunicación está provocando un cambio en profundidad en las profesiones i sectores productivos que tienen el lenguaje como herramienta de trabajo. La enseñanza, la traducción, la edición de textos y la elaboración de diccionarios y enciclopedias han experimentado cambios profundos, que exigen actualizar los conocimientos y los métodos utilizados por los profesionales de estas áreas.

Esta situación abre nuevas perspectivas profesionales a los filólogos y a los lingüistas y, a su vez, define un nuevo perfil de profesional: además del conocimiento lingüístico hay que adquirir conocimientos de informática y de lingüística computacional, tanto en el campo del procesamiento de la lengua escrita como de la lengua hablada, poniendo un énfasis especial en la capacidad de abstracción, el razonamiento lógico y la capacidad de organización y estructuración de los datos.

Diferentemente a lo que pasaba con la filología tradicional, el desarrollo de sistemas de procesamiento de la información, del lenguaje y del habla requiere la colaboración de especialistas de áreas diferentes. Por este motivo, en el perfil del lingüista que quiere centrar su actividad en el nuevo entorno de la sociedad de la información es esencial una buena capacidad para el trabajo en equipo y para la organización y gestión de proyectos multidisciplinares, donde cada una de las tareas individuales se orienta a la consecución de un objetivo común.



Bibliografía:

ABAITUA ODRIEZOLA, J.K. (1996) "Ingeniería de la lengua y normalización lingüística". En: FORCADA, V.M.; SAGER, J. C.; GIL DE CARRASCO, A. (Eds.) *Estudios computacionales del español y del inglés*. (pág. 253-265) Madrid: Instituto Cervantes.
<http://sirio.deusto.es/abaitua/konzeptu/umist.htm>

BLECUA, J.M.; CLAVERÍA, G.; SÀNCHEZ, C.; TORRUELLA, J. (Eds.) (1999) *Filología e informática. Nuevas tecnologías en los estudios filológicos*. Barcelona: Seminario de Filología e Informática, Departamento de Filología Española, Universidad Autónoma de Barcelona - Editorial Milenio.

COLE, R.A.; MARIANI, J.; USZKOREIT, H.; ZAENEN, A.; ZUE, V. (Eds.) (1997) *Survey of the State of the Art in Human Language Technology*. Cambridge: Cambridge University Press.
<http://cslu.cse.ogi.edu/HLTsurvey/HLTsurvey.html>

DANZIN, A. (1992) *Vers une infrastructure linguistique européenne*. Rapport présenté par A. Danzin et le Groupe de réflexion stratégique pour la CCE (DGXIII). Versión española: *Hacia una infraestructura lingüística europea*. Informe presentado por A. Danzin y el Grupo de reflexión estratégica para la Comisión de las Comunidades Europeas (D.G. XIII). 31.03.1992.

EDWARDS, J.; KINGSCOTT, G. (1996) *Language Industries Atlas*. Amsterdam: IOS Press.
<http://158.169.50.95:10080/mlis/en/atlas/home.html>

GÓMEZ GUINOVART, J. (1998) "Fundamentos de lingüística computacional: bases teóricas, líneas de investigación y aplicaciones". En: CID LEAL, P.; J. BARÓ y QUERALT (Eds.) *Anuario SOCADI de Documentación e Informació 1998*. (pág. 135-146) Barcelona: Societat Catalana de Documentació i Informació.

GÓMEZ GUINOVART, J.; LORENZO SUÀREZ, A.; PÉREZ GUERRA, J.; ÀLVAREZ LUGRÍS, A. (Eds.) (1999) *Panorama de la investigación en lingüística informática*. RESLA, *Revista Española de Lingüística Aplicada*, Volumen monográfico.

GÓMEZ GUINOVART, J.; PALOMAR, M. (Coords.) (1998) *Monografía: Lengua y Tecnologías de la Información*. *Novática, Revista de la Asociación de Técnicos de Informática*, 133 (Mayo-Junio). pág. 3-45.

Green Paper on Living and Working in the Information Society: People First. European Commission, Directorate General V, Employment, Industrial Relations and Social Affairs. COM (96) 389, 22.07.96.
<http://europa.eu.int/ISPO/infosoc/legreg/docs/peopl1st.html>

Human language technologies. Living and working together in the Information Society. Discussion document. Luxembourg, DG XIII/E-5, July 1997.
<http://www.hltcentral.org/htmlengine.shtml?id=221>

Human Language, Human Communication, Human Knowledge. Summary Report. Luxembourg, DG XIII/E-5, July 1998.
<http://www.hltcentral.org/htmlengine.shtml?id=222.0>

Ingeniería lingüística. Cómo aprovechar la fuerza del lenguaje. Luxemburg: Anite Systems. Versión española a cargo del Observatorio Español de Industrias de la Lengua, Instituto Cervantes.
http://www.hltcentral.org/usr_docs/project-source/en/broch/harness.html

Ingeniería lingüística: un mayor uso del lenguaje en todos los sentidos. Luxemburg: Cray Systems.
http://www.hltcentral.org/usr_docs/project-source/en/whatis.html

KURZWEIL, R. (1998) "When Will HAL Understand What We Are Saying? Computer Speech Recognition and Understanding". En: STORK, D.G. (Ed.) *Hal's Legacy: 2001's Computer es Dream and Reality*. Cambridge, Mass.: The MIT Press.



<http://mitpress.mit.edu/e-books/Hal/chap7/seven1.html>

L'ingénierie linguistique. Délégation Générale à la Langue Française, Réseau international des observatoires francophones de l'informate et du traitement informatique des langues.

<http://www.culture.fr/culture/dgff/rioifil/enjeux.htm>

MARTÍ ANTONIN, M.A. (1999) "Panorama de la lingüística computacional en Europa". En: GÓMEZ GUINOVART, J.; LORENZO SUÀREZ, A.; PÉREZ GUERRA, J.; ÀLVAREZ LUGRÍS, A. (Eds.) *Panorama de la investigación en lingüística informática*. RESLA, *Revista Española de Lingüística Aplicada*, Volumen monográfico. pág. 11-24.

MARTÍ ANTONIN, M.A. (Coord.) (1999) *Llengua catalana IV (Tecnologies de la llengua)*. Barcelona: Edicions de la UOC.

MARTÍ ANTONIN, M. A.; CASTELLÓN MASALLES, I. (2001) *La lingüística computacional*. Barcelona: Publicacions de la Universitat de Barcelona.

MORENO SANDOVAL, A. (1998) *Lingüística Computacional*. Madrid: Síntesis.

MOURE, T.; LLISTERRI, J. (1996) "Lenguaje y nuevas tecnologías. El campo de la lingüística computacional". En: FERNÁNDEZ PÉREZ, M. (Coord.) *Avances en lingüística aplicada*. Santiago de Compostela: Universidad de Santiago de Compostela, Servicio de Publicaciones e Intercambio Científico (Avances en...; 4). pág. 147-228.

OLIVE, J.P. (1997) "The Talking Computer Text to Speech Synthesis". En: STORK, D.G. (Ed.) *Hal's Legacy: 2001's Computer es Dream and Reality*. Cambridge, Mass.: The MIT Press.

<http://mitpress.mit.edu/e-books/Hal/chap6/six1.html>

Procesamiento del lenguaje natural. Número monográfico. *Philologia Hispalensis*. Vol. 11, núm. 2 (1997).

The EUROMAP Report. Challenge & Opportunity for Europe's Information Society. September 1998. DG XIII, Telematics Applications Programme.

http://www.hltcentral.org/usr_docs/Euromap-report/euromap97.zip

The Language Engineering Directory. Compiled by Language & Technology S.L. Madrid, 1996.

<http://158.169.50.95:10080/mlis/en/direct/home.html>

ZUE, V. (1999) "Talking with your computer", *Scientific American*, August 1999. pág. 40-41.

<http://www.sciam.com/1999/0899issue/0899zue.html>

Enlaces relacionados:

- ★ V Programa marco de investigación y desarrollo de la Unión Europea:
<http://www.cordis.lu/fp5/home.html>
- ★ Secretariado de la Sociedad de la Información (Generalitat de Catalunya):
http://www.gencat.es/csi/cat/societat_informacio/
- ★ Survey of the State of the Art in Human Language Technology:
<http://cslu.cse.ogi.edu/HLTsurvey/HLTsurvey.html>
- ★ Centro de Lenguaje y Computación (CLiC):
<http://clic.fil.ub.es>

[Fecha de publicación: mayo de 2001]



Cita recomendada:

MARTÍ ANTONIN, Maria Antònia; LLISTERRI, Joaquim (2001). "La ingeniería lingüística en la sociedad de la información". *Digithum*, n.º 3 [artículo en línea].
DOI: <http://dx.doi.org/10.7238/d.v0i3.595>