

Las tecnologías lingüísticas

Tecnologías lingüísticas, ingeniería lingüística e industrias de la lengua

Términos relacionados con las tecnologías lingüísticas

La distinción entre “texto” y “habla”

Ámbitos relacionados con las tecnologías lingüísticas

Los sistemas de procesamiento del lenguaje y del habla

El desarrollo de las aplicaciones

Tecnologías lingüísticas

Tecnologías lingüísticas, ingeniería lingüística e industrias de la lengua

Tecnologías lingüísticas

TL, Tecnologías lingüísticas, Tecnologías del lenguaje, Tecnologías de la lengua.

TL, *Tecnologies lingüístiques*, *Tecnologies del llenguatge*, *Tecnologies de la llengua*.

LT, *Language Technologies*.

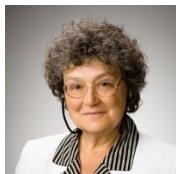
TLH, Tecnologías del lenguaje humano.

TLH, *Tecnologies del llenguatge humà*.

HLT, *Human Language Technologies*.

“El término tecnologías lingüísticas (TL) abarca el conjunto de los componentes, herramientas, técnicas y aplicaciones de software que procesan el lenguaje humano natural” (p. 5).

Lockwood, R., y Joscelyne, A. (2003). *Resumen del estudio comparativo de la evolución de las tecnologías lingüísticas en Europa*. Copenhague - Alcalá de Henares: EUROMAP Language Technologies, Center for Sprogteknologi - Oficina del Español en la Sociedad de la Información, Instituto Cervantes. Consultado en http://liceu.uab.cat/~joaquim/language_technology/docs/Lockwood_Joscelyne_03_Estudio_Comparativo_Euromap.pdf



Rose Lockwood Andrew Joscelyne

“Language technologies are technologies for automatically analysing and generating the most complex information medium in our world, human language, in both its spoken and written forms (as well as sign language). These technologies are developed by experts involved in linguistics, computer science, computational linguistics and related disciplines.”

Rehm, G. y Uszkoreit, H. (Eds.). (2013). *Strategic research agenda for multilingual Europe 2020*. Berlin - Heidelberg: Springer. Consultado en <http://link.springer.com/book/10.1007/978-3-642-36349-8/page/1>



Georg Rhem

Hans Uszkoreit

“Por tecnologías de la lengua o ingeniería lingüística se entiende los programas que procesan el lenguaje humano con los siguientes objetivos: mejorar la comunicación en todas sus modalidades y facilitar el acceso a la información por encima de las barreras que impone la distancia, el uso de lenguas distintas o el modo en que tiene lugar la comunicación, ya sea hablado o escrito . . . Se trata, en último término, de aplicar los conocimientos sobre la lengua al desarrollo de sistemas informáticos, con el fin de que puedan reconocer, comprender e interpretar y generar el lenguaje humano en todas sus formas” (p. 9).

Martí, M. A. (2003). Las tecnologías de la lengua y la sociedad de la información. En M. A. Martí (Ed.), *Tecnologías del lenguaje* (pp. 9-29). Barcelona: Editorial UOC.



Maria Antònia Martí

Suárez, A. (2014). Introducción a las TLH. *moocTLG. Nuevos retos en las Tecnologías del Lenguaje Humano*. Departamento de Lenguajes y Sistemas Informáticos, Universidad de Alicante. Consultado en <https://www.youtube.com/watch?t=419&v=DEG2odTevs>

Ingeniería lingüística

IL, Ingeniería lingüística.

EI, *Enginyeria lingüística*.

LI, *Language Engineering*.

“La ingeniería lingüística es la aplicación de los conocimientos sobre la lengua al desarrollo de sistemas informáticos para que puedan reconocer, comprender, interpretar y generar el lenguaje humano en todas sus formas” (p. 5).

Ingeniería lingüística. Cómo aprovechar la fuerza del lenguaje. Folleto preparado por Anite Systems para el proyecto LINGLINK en nombre de los participantes del sector «Ingeniería Lingüística» del Programa de Aplicaciones Telemáticas y adaptado al español. (s. f.). Luxembourg: Anite Systems. Consultado en http://liceu.uab.cat/~joaquim/language_technology/docs/Como_aprovechar_la_fuerza_del_lenguaje.pdf

“La ingeniería lingüística es una tecnología que utiliza nuestros conocimientos de la lengua para mejorar la utilización de los sistemas informáticos, de manera que:

- perfecciona nuestra interacción con ellos;
- asimila, analiza, selecciona, y utiliza y presenta la información de manera más eficaz;
- proporciona medios de generación y de traducción del lenguaje natural” (p. 5).

Ingeniería lingüística. Cómo aprovechar la fuerza del lenguaje. Folleto preparado por Anite Systems para el proyecto LINGLINK en nombre de los

participantes del sector «Ingeniería Lingüística» del Programa de Aplicaciones Telemáticas y adaptado al español. (s. f.). Luxembourg: Anite Systems. Consultado en http://liceu.uab.cat/~joaquim/language_technology/docs/Como_aprovechar_la_fuerza_del_lenguaje.pdf

Industrias de la lengua

“Ce terme désigne l’ensemble des activités qui visent à faire manipuler, interpréter ou générer par des machines le langage naturel écrit ou parlé par les humains” (p. 8).

Carré, R., Dégremon, J. F., Gross, M., Pierrel, J. M. y Sabah, G. (1991). *Langage humain et machine*. Paris: Presses du CNRS.

“La industria de la lengua . . . abarca una serie de actividades comerciales en las que el tratamiento del lenguaje, por personas, por máquinas o por una combinación de unas y otras, forma una parte fundamental del producto o servicio. De este modo, la industria de la lengua incluye métodos de publicación nuevos y tradicionales, desde los libros a las páginas de la *World Wide Web*, desde la radio a la televisión de pago. Abarca servicios ya consolidados, como la traducción, la interpretación, la transcripción y la redacción técnica, y otros servicios nuevos, como la adaptación de programas informáticos, el correo vocal y el aprendizaje de idiomas asistido por ordenador. Incluye productos con una larga historia como la pluma y el lápiz, y una amplia variedad de productos informáticos para el tratamiento de las lenguas naturales, como máquinas de bolsillo de traducción de frases, procesadores de textos, bases de datos terminológicas y sistemas de traducción automática” (p. 12).

Lenguaje y tecnología. De la torre de Babel a la aldea global. (1997). Luxemburgo: Oficina de Publicaciones Oficiales de las Comunidades Europeas. Consultado en http://liceu.uab.cat/~joaquim/language_technology/docs/Lenguaje_Tecnologia_MLIS_97.pdf

“toutes les activités techniques et économiques liées à la mise à la disposition de ces nouveaux outils du langage: ressources linguistiques de base, moyens de conversion entre le langage parlé et le langage écrit, traitement sémantiques et syntaxiques, passages d’une langue à l’autre, analyse de textes et analyses d’élocution et toutes autres interventions par voie technique sur l’usage des langues naturelles. Plus généralement, tout ce qui modifie la langue écrite et parlée par l’usage des nouvelles technologies quelles qu’elles soient.”

Danzin, A. (1992). *Vers une infrastructure linguistique européenne*. Rapport présenté par A. Danzin et le Groupe de Réflexion Stratégique pour la Commission des Communautés Européennes (DG XIII).



Términos relacionados con las tecnologías lingüísticas

Lingüística computacional

LC, Lingüística computacional.

LC, *Lingüística computacional*.

CL, *Computational Linguistics*.

“*Computational linguistics* is the scientific study of language from a computational perspective. Computational linguists are interested in providing computational models of various kinds of linguistic phenomena. These models may be “knowledge-based” (“hand-crafted”) or “data-driven” (“statistical” or “empirical”). Work in computational linguistics is in some cases motivated from a scientific perspective in that one is trying to provide a computational explanation for a particular linguistic or psycholinguistic phenomenon; and in other cases the motivation may be more purely technological in that one wants to provide a working component of a speech or natural language system. Indeed, the work of computational linguists is incorporated into many working systems today, including speech recognition systems, text-to-speech synthesizers, automated voice response systems, web search engines, text editors, language instruction materials, to name just a few.”

Association for Computational Linguistics. (s. f.). What is computational linguistics? *ACL Homepage*. Stroudsburg PA: Association for Computational Linguistics. Consultado en <http://aclweb.org/website/what-is-cl>

Oliver, A. (2014). Basado en corpus y basado en conocimiento. *moocTLG. Nuevos retos en las Tecnologías del Lenguaje Humano*. Departamento de Lenguajes y Sistemas Informáticos, Universidad de Alicante. Consultado en <https://www.youtube.com/watch?t=130&v=nW5P0xye4pI>

“La LC se ha centrado en la investigación básica sobre el análisis del lenguaje, los formalismos de representación de la información lingüística, la investigación sobre técnicas de análisis y el desarrollo de lenguajes de representación y adquisición del conocimiento . . . siendo su objetivo último la modelización del conocimiento lingüístico para hacer posible la construcción de sistemas computacionales que comprendan y generen textos en lenguaje natural” (p. 10).

Martí, M. A. (2003). Las tecnologías de la lengua y la sociedad de la información. En M. A. Martí (Ed.), *Tecnologías del lenguaje* (pp. 9-29). Barcelona: Editorial UOC.



Maria Antònia Martí

“La Lingüística Computacional es un área interdisciplinaria entre la Lingüística y la Informática que se ocupa de la construcción de sistemas informáticos capaces de procesar el lenguaje humano” (p. 73).

“En su vertiente *teórica*, la LC se centra en una serie de investigaciones que se ocupan de la formalización del conocimiento lingüístico que los humanos necesitamos para generar y comprender el lenguaje. . . . En su vertiente *aplicada*, la LC se centra en los resultados prácticos de los modelos del uso del lenguaje. Los métodos, técnicas, herramientas y aplicaciones en esta área se subsumen con frecuencia bajo el término de *Ingeniería lingüística* o *Tecnologías del Lenguaje*” (pp. 76-77).

Lavid, J. (2005). *Lenguaje y nuevas tecnologías. Nuevas perspectivas, métodos y herramientas para el lingüista del siglo XXI*. Madrid: Cátedra.



Julia Lavid

Procesamiento del lenguaje natural

PLN, Procesamiento del lenguaje natural.

PLN, *Processament del llenguatge natural*.

NLP, *Natural Language Processing*.

TALN, *Traitement automatique du langage naturel*.

“se centra en los aspectos más aplicados de la LC, ya que trata de buscar soluciones a los problemas que plantea la

comprensión del lenguaje natural en el marco de sistemas concretos . . . En general, las técnicas de procesamiento del lenguaje natural están orientadas a la construcción de una representación del contenido de las producciones lingüísticas en términos de un determinado metalenguaje, y el núcleo de interés de las mismas lo constituyen el desarrollo de gramáticas y analizadores” (p. 10).

Martí, M. A. (2003). Las tecnologías de la lengua y la sociedad de la información. En M. A. Martí (Ed.), *Tecnologías del lenguaje* (pp. 9-29). Barcelona: Editorial UOC.



Maria Antònia Martí

“L’objectif des traitements automatiques des langues est la conception de logiciels (programmes) capables de traiter de façon automatique des données linguistiques, c’est à dire des données exprimées dans une langue (dite naturelle)” (p. 27).

Fuchs, C. (Ed). (1993). *Linguistique et traitement automatique des langues*. Paris: Hachette.



Catherine Fuchs

Tecnologías del habla

El ámbito de las tecnologías del habla

Lingüística informática

“une branche de la linguistique formelle qui recourt à l’ordinateur comme outil de validation d’hypothèses théoriques sur le fonctionnement du langage” (p. 22).

Fuchs, C. (Ed). (1993). *Linguistique et traitement automatique des langues*. Paris: Hachette



Catherine Fuchs

“orientada al desarrollo de programas de apoyo a los estudios filológicos, lexicográficos, lingüísticos, humanísticos, etc.” (p. 11).

Martí, M. A. (2003). Las tecnologías de la lengua y la sociedad de la información. En M. A. Martí (Ed.), *Tecnologías del lenguaje* (pp. 9-29). Barcelona: Editorial UOC.



Maria Antònia Martí

Informática lingüística

“conçue dans son principe comme une branche de l’informatique, qui ne recourt qu’aux méthodes et outils de l’informatique, et dont le domaine d’application se trouve être des données linguistiques . . . prise dans un sens large, l’expression en vient à

désigner l'ensemble des traitements automatiques de données linguistiques" (p. 22).

Fuchs, C. (Ed). (1993). *Linguistique et traitement automatique des langues*. Paris: Hachette



Catherine Fuchs



La división entre “lenguaje” y “habla”

Dos comunidades bien diferenciadas en el ámbito de las tecnologías del lenguaje en función de la modalidad lingüística considerada: oral o escrita.

Tratamiento de la lengua escrita

Expertos dedicados al procesamiento del lenguaje natural o a la lingüística computacional: *the natural language (processing) community*, *the NLP community*.



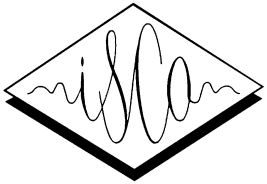
ACL, *Association for Computational Linguistics*



SEPLN, *Sociedad Española para el Procesamiento del Lenguaje Natural*

Tratamiento de la lengua oral

Expertos dedicados a las tecnologías del habla: *the speech community*.



ISCA, *International Speech Communication Association*



RTTH, *Red Temática en Tecnologías del Habla*

Tratamiento de la lengua escrita y de la lengua oral

TLH, *Tecnologías del lenguaje y del habla*.

TLP, *Tecnologies del llenguatge i de la parla*.

SLT, *Speech and Language Technologies*.

Denominación integradora que abarca tanto el tratamiento de la lengua escrita como el de la lengua hablada.



Ámbitos relacionados con las tecnologías lingüísticas

Las tecnologías lingüísticas constituyen un campo interdisciplinario.

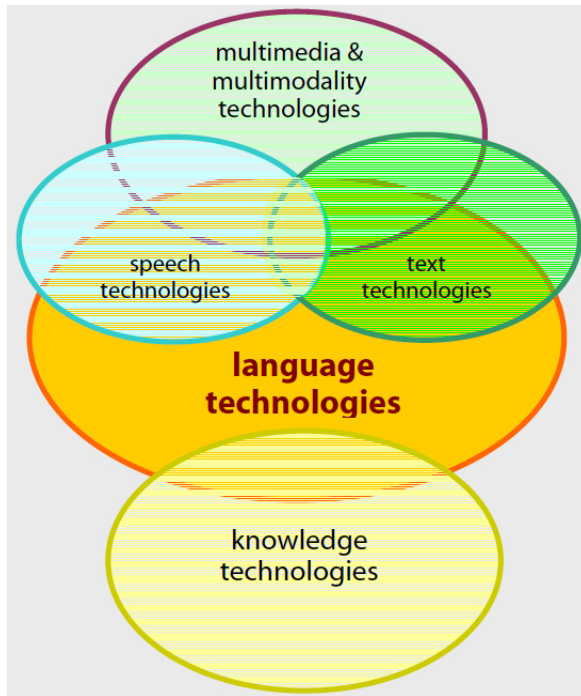
“complex information and most of human knowledge is maintained and transmitted in written texts. Speech and text technologies process or produce language in these two modes of realization. But language also has aspects that are shared between speech and text such as dictionaries, most of grammar and the meaning of sentences. Thus large parts of language technology cannot be subsumed under speech and text technologies. Among those are technologies that link language to knowledge. We do not know how language, knowledge and thought are represented in the human brain. Nevertheless, language technology had to create formal representation systems that link language to concepts and tasks in the real world. This provides the interface to the fast growing area of knowledge technologies.

In our communication we mix language with other modes of communication and other information media. We combine speech with gesture and facial expressions. Digital texts are combined with pictures and sounds. Movies may contain language and spoken and written form. Thus speech and text technologies overlap and interact with many other technologies that facilitate processing of multimodal communication and multimedia documents.”

Uszkoreit, H. (s.d.). *Language technology. A first overview*. Department of Computational Linguistics and Phonetics, Saarland University at Saarbrücken. Consultado en <http://www.dfki.de/%7Ehansu/LT.pdf>



Hans Uszkoreit



Las tecnologías lingüísticas.

Uszkoreit, H. (s.d.). *Language technology. A first overview*. Department of Computational Linguistics and Phonetics, Saarland University at Saarbrücken. Consultado en <http://www.dfki.de/%7Ehansu/LT.pdf>

Lingüística

Ofrece una descripción y una modelización del lenguaje (humano) y de las lenguas (naturales).

Informática

Ciencias de la computación.

Ciències de la computació.

CS, *Computer Sciences*.

Proporciona teorías, herramientas y técnicas para el tratamiento automático de la lengua.

Inteligencia artificial

IA, Inteligencia artificial.
AI, Intel·ligència artificial.
AI, *Artificial intelligence*.

Desarrollo de sistemas informáticos que simulan el comportamiento humano en lo que se refiere a:

- adquisición de conocimientos;
- razonamiento a partir de los conocimientos adquiridos;
- representación de los conocimientos adquiridos.

Ingeniería de telecomunicación

Centrada en la transmisión y el tratamiento del habla en tanto que señal sonora.

Psicología

Proporciona una descripción y una modelización de la interacción y del comportamiento lingüístico humano.

- Procesos cognitivos.
- Factores humanos.
- Usabilidad.



Los sistemas de procesamiento del lenguaje y del habla

Un sistema de tratamiento automático del lenguaje o del habla implica una serie de procesos:

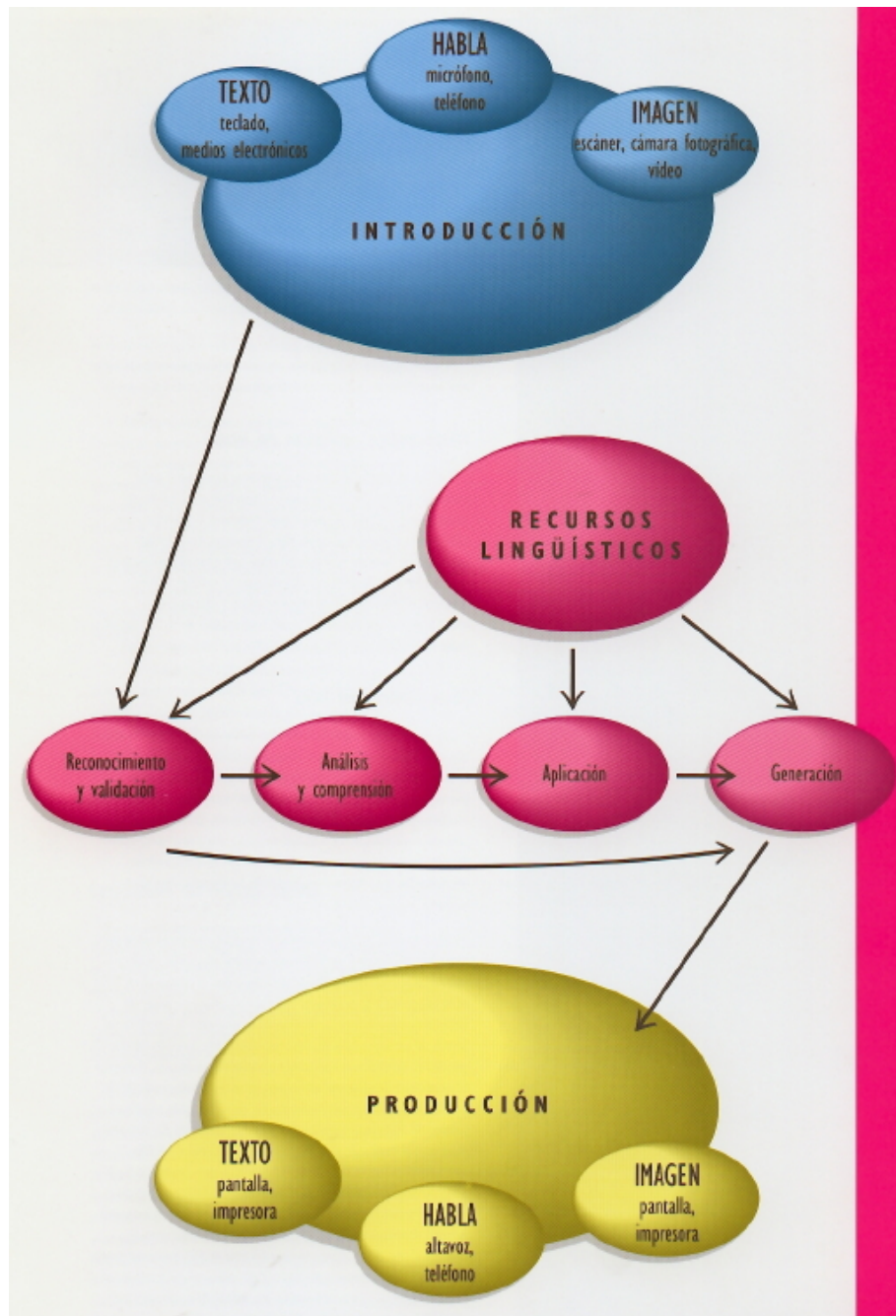
- Adquisición de los datos en formato digital.
- Codificación del documento o de la señal sonora.

Tratamiento o procesamiento de los datos: [conversión de texto en habla](#), [conversión de habla en texto](#), [recuperación](#) y

[extracción](#) de información, [traducción de texto](#) o [traducción de habla...](#)

Creación de un nuevo documento o de una señal sonora con el resultado del tratamiento.

Presentación de los resultados del tratamiento.



Modelo de un sistema con capacidad lingüística.

Ingeniería lingüística. Cómo aprovechar la fuerza del lenguaje. Folleto preparado por Anite Systems para el proyecto LINGLINK en nombre de los participantes del sector «Ingeniería Lingüística» del Programa de Aplicaciones Telemáticas y adaptado al español. (s. f.). Luxembourg: Anite Systems. Consultado en http://liceu.uab.cat/~joaquim/language_technology/docs/Como_aprovechar_la_fuerza_del_lenguaje.pdf (p. 6).



El desarrollo de aplicaciones

Investigación y desarrollo.

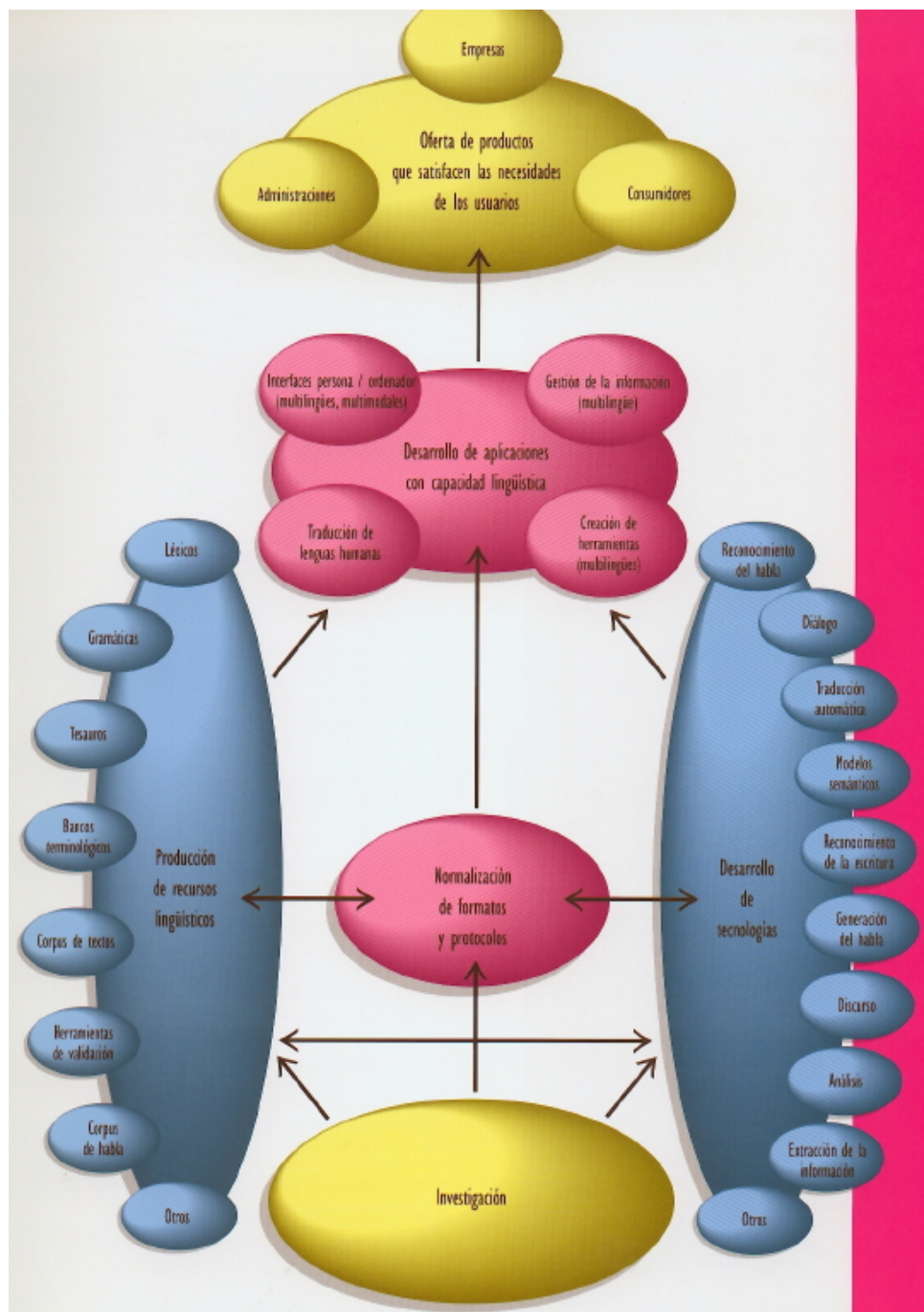
Tecnologías lingüísticas.

Recursos lingüísticos.

Estándares.

Aplicaciones que integran funciones lingüísticas.

Servicios y productos para el usuario.



Modelo de actividades de ingeniería lingüística.

Ingeniería lingüística. Cómo aprovechar la fuerza del lenguaje. Folleto preparado por Anite Systems para el proyecto LINGLINK en nombre de los participantes del sector «Ingeniería Lingüística» del Programa de Aplicaciones Telemáticas y adaptado al español. (s. f.). Luxembourg: Anite Systems. Consultado en http://liceu.uab.cat/~joaquim/language_technology/docs/Como_aprovechar_la_fuerza_del_lenguaje.pdf (p. 12).



Las tecnologías lingüísticas

[Joaquim Llisterri](#), Departament de Filologia Espanyola, Universitat Autònoma de Barcelona

http://liceu.uab.cat/~joaquim/language_technology/HLT/tecnol_ling_gen.html

Last updated: 29/10/15 08:29



This work is licensed under a [Creative Commons Attribution-Noncommercial-Share Alike 3.0 License](#).