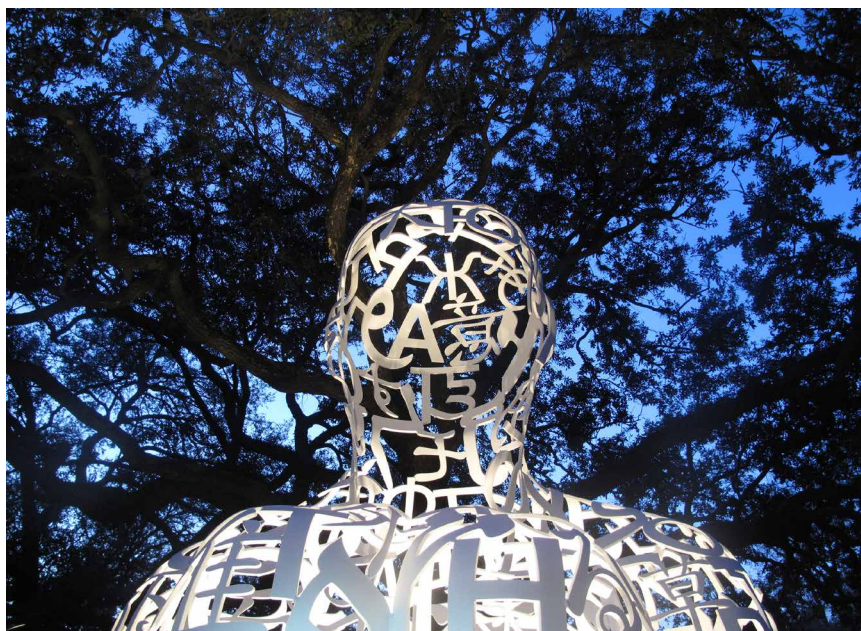


Terminología: entre lo humano y lo tecnológico



Elea Giménez Toledo
Elena Montiel-Ponsoda
Nava Maroto
(eds.)

Prólogo de M. Teresa Cabré

EDITORIAL COMARES



Interlingua

TERMINOLOGÍA: ENTRE LO HUMANO Y LO TECNOLÓGICO

Elea Giménez Toledo
Elena Montiel-Ponsoda
Nava Maroto
(eds.)

Terminología: entre lo humano y lo tecnológico

Prólogo de
M. Teresa Cabré

Granada, 2026

Colección indexada en la MLA International Bibliography desde 2005

EDITORIAL COMARES

INTERLINGUA

427

Colección fundada por:

EMILIO ORTEGA ARJONILLA Y PEDRO SAN GINÉS AGUILAR

Comité Científico (Asesor):

ESPERANZA ALARCÓN NAVÍO Universidad de Granada	CATALINA JIMÉNEZ HURTADO Universidad de Granada
JESÚS BAIGORRI JALÓN Universidad de Salamanca	ÁNGELA LARREA ESPINAR Universidad de Córdoba
CHRISTIAN BALLIU ISTI, Bruxelles	HELENA LOZANO Università di Trieste
LORENZO BLINI LUSPIO, Roma	JAVIER MARTÍN PÁRRAGA Universidad de Córdoba
ANABEL BORJA ALBÍ Universitat Jaume I de Castellón	ANTONIO RAIGÓN RODRÍGUEZ Universidad de Córdoba
NICOLÁS A. CAMPOS PLAZA Universidad de Murcia	MARIA JOÃO MARÇALO Universidade de Évora
MIGUEL Á. CANDEL-MORA Universitat Politècnica de València	FRANCISCO MATTE BON LUSPIO, Roma
ÁNGELA COLLADOS AÍS Universidad de Granada	CHELO VARGAS-SIERRA Universidad de Alicante
MIGUEL DURO MORENO Universidad de Málaga	MERCEDES VELLA RAMÍREZ Universidad de Córdoba
FRANCISCO J. GARCÍA MARCOS Universidad de Almería	ÁFRICA VIDAL CLARAMONTE Universidad de Salamanca
GLORIA GUERRERO RAMOS Universidad de Málaga	GERD WOTJAK Universidad de Leipzig

ENVÍO DE PROPUESTAS DE PUBLICACIÓN:

Las propuestas de publicación han de ser remitidas (en archivo adjunto, con formato PDF) a alguna de las siguientes direcciones electrónicas: anabelen.martinez@uco.es, psgines@ugr.es

Antes de aceptar una obra para su publicación en la colección INTERLINGUA, esta habrá de ser sometida a una revisión anónima por pares. Para llevarla a cabo se contará, inicialmente, con los miembros del comité científico asesor. En casos justificados, se acudirá a otros especialistas de reconocido prestigio en la materia objeto de consideración.

Los autores conocerán el resultado de la evaluación previa en un plazo no superior a 60 días. Una vez aceptada la obra para su publicación en INTERLINGUA (o integradas las modificaciones que se hiciesen constar en el resultado de la evaluación), habrán de dirigirse a la Editorial Comares para iniciar el proceso de edición.

Los resultados de investigación que aquí se presentan se enmarcan en el proyecto de investigación «TeresIA: terminología e inteligencia artificial», financiado con fondos de la Unión Europea Next Generation EU, Plan de Recuperación, Transformación y Resiliencia (PRTR), a través de la Secretaría de Estado de Digitalización e Inteligencia Artificial (SEDIA), Ministerio para la Transformación Digital y de la Función Pública.



© Imagen de portada: Jaume Plensa. *Mirror*, 2011. Acero inoxidable pintado, 377x253x245 cm. Rice University, Houston, Texas, USA. (A gift from Bill and Stephanie Sicks). Foto: Laura.Medina@PlensaStudioBarcelona.com.

© Los autores

© Editorial Comares, 2026

Polígono Juncaril • C/ Baza, parcela 208 • 18220 Albolote (Granada) España • Tlf.: 958 465 382

www.comares.com • E-mail: libreriacomares@comares.com

facebook.com/Comares • twitter.com/comareseditor • instagram.com/editorialcomares

ISBN: 979-13-7033-126-9 • Depósito legal: Gr. 763/2026

Impresión y encuadernación: COMARES

Contenido

Prólogo. DE TERMINESP A TERESIA, UN PROYECTO CAPITAL PARA LA TERMINOLOGÍA DEL ESPAÑOL	IX
<i>M. Teresa Cabré</i>	
INTRODUCCIÓN	1
<i>Elea Giménez Toledo, Elena Montiel-Ponsoda y Nava Maroto</i>	
LA CONFIGURACIÓN DE LA TERMINOLOGÍA DEL ESPAÑOL	
1. LA CONFIGURACIÓN DE LA TERMINOLOGÍA DEL ESPAÑOL: DE LOS ORÍGENES A TERESIA.	9
<i>Joaquín García Palacios y Cecilio Garriga Escribano</i>	
TERMINOLOGÍA, CIENCIA ABIERTA Y MULTILINGÜISMO EN LA CIENCIA	
2. LA IMPRONTA DE LA POLÍTICA CIENTÍFICA EUROPEA EN TERESIA: INTERDISCIPLINARIEDAD, CIENCIA ABIERTA, MULTILINGÜISMO E IMPACTO SOCIAL	29
<i>Elea Giménez Toledo</i>	
3. DATOS ABIERTOS Y ENLAZADOS EN TERMINOLOGÍA	39
<i>Elena Montiel-Ponsoda y Patricia Martín-Chozas</i>	
UNIDADES TERMINOLÓGICAS: CONOCIMIENTO EXPERTO Y TECNOLOGÍA PARA SU DETECCIÓN Y ESTUDIO	
4. MÉTODOS DE EXTRACCIÓN AUTOMÁTICA DE TÉRMINOS Y RELACIONES EN EL DOMINIO JURÍDICO	59
<i>Patricia Martín-Chozas, Pablo Calleja, Joel Pardo-Ferrera y Elena Montiel-Ponsoda</i>	
5. IMPLEMENTACIÓN DE SISTEMAS DE EXTRACCIÓN AUTOMÁTICA DE TÉRMINOS CLAVE MEDIANTE TECNOLOGÍAS DEL LENGUAJE E INTELIGENCIA ARTIFICIAL	79
<i>Sergi Marsol Torrent, Salvador Lima-López y Martin Krallinger</i>	
6. «CORPUS TERESIA»: UN CORPUS ANOTADO DE TERMINOLOGÍA BIOMÉDICA Y JURÍDICA. DISEÑO DE GUÍAS, PROCESO DE ANOTACIÓN Y VALIDACIÓN METODOLÓGICA	97
<i>Victoria Saura-Montesinos, Salvador Lima-López, Judith Rosell, Elena Montiel-Ponsoda y Martin Krallinger</i>	

TERMINOLOGÍA: ENTRE LO HUMANO Y LO TECNOLÓGICO

7. ANÁLISIS DE LAS UNIDADES TERMINOLÓGICAS ANOTADAS EN EL CORPUS DE REFERENCIA TERESIA. . . 119
Rosa Castro-Prieto, Arsenio Andrades y Nava Maroto

CREANDO LAZOS ENTRE TERMINOLOGÍAS

8. REPRESENTACIÓN DE TERMINOLOGÍAS EN FORMATOS INTEROPERABLES. 143
Paula Díez-Ibarbia, Patricia Martín-Chozas y Elena Montiel-Ponsoda
9. DESAMBIGUACIÓN PARA EL ENLAZADO TERMINOLÓGICO. CORRESPONDENCIAS CONCEPTUALES ENTRE
RECURSOS TERMINOLÓGICOS DE ACUERDO AL PARADIGMA DE LOS DATOS ENLAZADOS 159
Lucía Palacios, Patricia Martín-Chozas, Elena Montiel-Ponsoda y Pablo Calleja

LO HUMANO Y LO TECNOLÓGICO EN LA VALIDACIÓN POR EXPERTOS Y EN LA SANCIÓN LINGÜÍSTICA

10. DE LAS PALABRAS A LAS MÁQUINAS: PARAMETRIZACIÓN DE CRITERIOS LINGÜÍSTICOS PARA EL DISEÑO
DE UN PROTOCOLO SEMIAUTOMATIZADO DE VALIDACIÓN DE NEOLOGISMOS ESPECIALIZADOS DEL
ESPAÑOL 175
Miguel Sánchez Ibáñez
11. TERMONIA: UN DESARROLLO TECNOLÓGICO PARA LA VALIDACIÓN Y LA SANCIÓN LINGÜÍSTICA DE
LOS TÉRMINOS 187
Adonis García Anés, Beatriz Guerrero García y Mariano Rico

UN LUGAR PARA TODAS LAS TERMINOLOGÍAS: EL METABUSCADOR TERESIA

12. CONSTRUCCIÓN DE UN METABUSCADOR PARA TERMINOLOGÍAS FAIR Y ENLAZADAS. ADAPTACIÓN
DE ONTOPORTAL A TERESIA 199
Carlos Badenes-Olmedo Y Markus Schüller Perdigon
13. HILVANANDO TERESIA: DEL TEXTO AL METABUSCADOR. SIMULACRO DE VALIDACIÓN Y SANCIÓN DE
TÉRMINOS BIOSANITARIOS EXTRAÍDOS DEL PROCESO DE ANOTACIÓN Y EVALUACIÓN PARA ENRIQUE-
CIMIENTO DE RECURSOS TERMINOLÓGICOS 211
Beatriz Guerrero García y Andrea Fernández Vivanco

APLICACIONES E IMPACTO SOCIAL DE LA TERMINOLOGÍA

14. TRADUCCIÓN INSTITUCIONAL Y TERESIA: GOBERNANZA TERMINOLÓGICA EN CONTEXTOS MULTI-
LINGÜES 227
Nava Maroto
15. LA CONTRIBUCIÓN DE TERESIA A LA CLARIFICACIÓN DEL DISCURSO TRIBUTARIO EN ESPAÑOL: IDEN-
TIFICACIÓN, CLASIFICACIÓN Y ANÁLISIS DE TERMINOLOGÍA ESPECIALIZADA 245
Mar Forment Fernández y Víctor González Ruiz

- AUTORES 263

Prólogo

De TERMINESP a TeresIA, un proyecto capital para la terminología del español

La terminología es una necesidad evidente para las lenguas de las sociedades desarrolladas. Mejor dicho es una necesidad evidente para todas las lenguas vivas. La razón es clara: si toda sociedad hablante se aproxima a la realidad y la proyecta verbalmente a través de las lenguas, y si esta sociedad distingue entre usos distintos de acuerdo con lo que desea comunicar y cómo desea comunicarlo, cuando necesita denominar algo muy preciso que no se confunda con otra cosa puede usar una combinación sintáctica (por ejemplo, un determinante definido) o un vocablo preciso. Es en esta necesidad de precisión que generalmente surgen los términos especializados.

En las sociedades avanzadas económicamente y productoras de conocimiento nuevo, la aparición de voces precisas para denominar los nuevos conceptos es un hecho natural, paralelo a la creación de conocimiento. Pero si además estas sociedades se relacionan económicamente con otras que manejan lenguas distintas, la terminología es imprescindible para asegurar una comunicación eficaz y segura en cuanto al sentido.

Una lengua considerada «lengua de cultura» debe facilitar a sus hablantes expresar y comunicar cualquier conocimiento en cualquier circunstancia y sobre cualquier temática, y además en distintos registros de comunicación. La transferencia de conocimiento especializado necesita precisión y, por tanto, no puede eludir el uso de terminología.

En el ámbito de la comunicación científico-técnica, así como en el ámbito de cualquier comunicación especializada en general, sea del comercio, de la industria, del deporte o de las finanzas, la terminología es absolutamente necesaria para que una lengua sea un instrumento seguro y completo de comunicación.

Pero ¿cómo surge la terminología en las lenguas?

En primer lugar, por la vía del uso. Si una lengua no se utiliza para hablar de ciertos temas, la terminología no aparece en el discurso sobre estos temas. El modo

natural de producción de terminología es que los hablantes de una lengua, que además son expertos en un ámbito de conocimiento, la utilicen en sus observaciones expertas innovadoras y en su transferencia a otros expertos o para transmitir conocimiento en contextos de formación de nuevos expertos.

La segunda vía por la que una lengua dispone de terminología propia es a través de la traducción y la interpretación. La falta de innovación en una comunidad hablante que, sin embargo, introduce conocimiento nuevo producido en otras sociedades y desea utilizarlo para su propia investigación o transmitirlo, puede generar también terminología. La importación de conocimiento nuevo original mediante la traducción es un hecho derivado directamente de la necesidad del traductor de buscar equivalentes para las unidades del texto que se va a traducir, en caso de no encontrarlos, debe decidir si crear un neologismo o adoptar el préstamo o adaptarlo. Pero parece evidente que adoptar préstamos de manera sistemática en un ámbito de especialidad no es la mejor estrategia para la conservación de una lengua.

Como hemos dicho, las lenguas de cultura científica usadas activamente para representar la especialidad requieren necesariamente terminología para convertirse en un instrumento de comunicación, pero para que esta terminología sea adecuada al registro funcional de la comunicación científica (del nivel más experto al nivel divulgativo) debe ser eficaz y contribuir a la precisión y a la concisión de este tipo de comunicación y, además, debe actualizarse permanentemente para poder participar del proceso de avance del conocimiento especializado. Esto requiere estar en continua renovación mediante la introducción de nuevos términos.

Dando un paso más allá, permítanme que ponga énfasis en otro aspecto de la terminología, en mi opinión necesario para que una lengua se sienta consolidada y segura: el reconocimiento de los términos como propios. Una lengua debe disponer de «terminología propia». Entendamos bien esta expresión: disponer de terminología significa que sus hablantes, en este caso los expertos de un ámbito de conocimiento, la sientan «suya» porque han participado implícita o explícitamente de su incorporación a la lengua. Implícitamente, porque asumen la política lingüística de actualización de esta lengua. Explícitamente, porque han consensuado los términos a través de su uso consciente. Los préstamos, de esta forma, devienen términos adoptados, con o sin adaptación, de acuerdo con las orientaciones de política de dicha lengua. Son términos aceptados y metabolizados en el discurso. Sin embargo, un ámbito de comunicación especializada que se actualiza sistemáticamente por la vía del préstamo corre el riesgo de pasar a ser a la larga «ámbito de comunicación perdido».

Ahora bien, las actitudes positivas en defensa de la calidad de una lengua solo se producen, como ya hemos dicho, cuando hay conciencia lingüística, conciencia que se manifiesta a través de la actitud de los propios especialistas o mediante procesos de normalización, consenso o armonización, ya sea regulados por los expertos desde el interior del ámbito, ya sea a través de una institución. Porque, aunque una lengua sea el

conjunto de sus usos reales, dejarla al arbitrio de su evolución exclusivamente natural puede conducir a la pérdida de su singularidad y al abandono de ámbitos de uso.

Esta regulación de diferente signo de la que acabamos de hablar es especialmente importante en lenguas que, como el español, son territorialmente muy extensas. En estas lenguas, si quiere preservarse la unidad por encima de las variedades, es evidente que hay que proceder a un proceso de armonización abierto y flexible acordado con los expertos. Esta validación da seguridad a los usuarios de la terminología.

Los hablantes de muchas de las lenguas minoritarias o minorizadas son conscientes, por su condición, de la necesidad de disponer de términos adecuados para poder producir comunicación en todos los ámbitos de especialidad, y la mayoría de estas lenguas disponen de políticas lingüísticas y de centros de terminología; pero en las lenguas consideradas mayoritarias, cuyo uso está asegurado por el número de hablantes, no se suele sentir esa necesidad. El español, por ejemplo, no dispone de una política explícita de armonización terminológica ni de centro alguno que se ocupe de este tema. Fue en el contexto de esta necesidad donde surgió la idea del proyecto TERMINESP, que describiremos muy sucintamente a continuación.

El proyecto TERMINESP, vinculado a la traducción, la normalización y la presencia internacional del español y del resto de lenguas de España, se inició en el año 2000 por iniciativa de la Asociación Española de Terminología (AETER), con la colaboración de otros organismos, con la finalidad de organizar, en una primera fase, la terminología del español. Por una razón evidente: los traductores, sobre todo los especializados y particularmente los institucionales, necesitaban terminología validada que asegurara la calidad de una traducción.

En este inicio constituyeron la comisión promotora: la Asociación Española de Terminología (AETER), representada por su presidenta; la Unidad de multilingüismo y coordinación terminológica de la Dirección General de la Traducción de la Comisión Europea; el Instituto Cervantes a través de la Unidad del Cervantes Virtual; la Fundación El español, lengua de traducción; la Red Iberoamericana de Terminología (RITERM); y la organización internacional Unión Latina.

No se trataba de proponer un mero proyecto de investigación, sino la creación de una infraestructura de organización de la terminología del español de España y su coordinación con la de los distintos países hispanohablantes. También se proponía enlazar la terminología del español con la de las lenguas cooficiales del Estado, que ya disponían de centros de terminología institucionales que coordinaban la producción, normalización y difusión de la terminología propia y ofrecían bases de datos terminológicos plurilingües, con inclusión del español. En julio de 2007 se incorporó el Consejo Superior de Investigaciones Científicas (CSIC).

La Asociación Española de Terminología (AETER) había nacido en el año 2003 por iniciativa de un grupo de especialistas e instituciones interesados, entre los que se contaban la Real Academia de Ciencias Exactas, Físicas y Naturales, el Consejo Superior de

Investigaciones Científicas, la Real Academia Española y los centros de terminología de las comunidades autónomas catalana (TERMCAT), vasca (UZEI) y gallega (TERMIGAL).

La necesidad de organizar armónicamente la creación de nuevos términos especializados, y particularmente los científicos y técnicos, se sentía como algo indispensable en comunidades preocupadas por el futuro de las lenguas propias, en situación minorizada. Esta preocupación surgió más tarde en comunidades lingüísticas afianzadas por el número de hablantes, pero interesadas por la calidad de la lengua en uso. Este fue el caso de la comunidad hispanohablante en su conjunto, para la que la terminología se consideró como una parte importantísima de su futuro, como afirmó el entonces rey de España don Juan Carlos I en el I Congreso Internacional de la Lengua Española que se llevó a cabo en la ciudad mexicana de Zacatecas en el año 2001.

En este contexto fue donde, desde AETER, se diseñó TERMINESP, con el objetivo de implementar un sistema de acceso y distribución en línea de los datos terminológicos del español, elaborados por muchos organismos, instituciones, empresas y particulares; de crear una comisión lingüística para la terminología del español que estableciera los criterios que debían orientar la creación, adaptación y adopción de nuevos términos; e interesar en su desarrollo a las instituciones más representativas de la lengua española, de la ciencia y la técnica, de la edición y de los medios de comunicación.

Los objetivos que TERMINESP proponía eran los siguientes:

- Organizar la terminología del español de España.
- Articular la organización de la terminología del español con la terminología de las distintas autonomías de lengua diferente al castellano: catalán, vasco y gallego.
- Impulsar la organización terminológica de los países hispanohablantes, especialmente los de América Latina.
- Y organizar una red que conjugara los nodos hispanoamericanos y los españoles peninsulares en una única organización.

Es evidente que se trataba de unos objetivos demasiado ambiciosos para una asociación como AETER. Pero, aun así, TERMINESP empezó a avanzar y a producir algunos resultados:

- Se creó una plataforma de almacenamiento de datos y se establecieron las vías de acceso a la terminología almacenada.
- Se extrajo semiautomáticamente toda la terminología de las Normas UNE de AENOR y se incorporó a la base de datos.
- Se hizo un catálogo de la terminología disponible en las universidades y centros de trabajo.
- Se planificó un proceso de validación de la terminología del español (VALITER), plataforma gestionada por la Unidad de Traducción de la Comisión Europea.
- Se organizó y se puso en marcha una comisión lingüística (COLTE), presidida por un miembro de la Real Academia de la Lengua Española (RAE) y cuya secre-

taría residía en la persona que ocupaba la presidencia de AETER, con la finalidad de establecer criterios lingüísticos para la terminología del español. Esta comisión redactó sus primeros documentos sobre formación de términos y criterios sobre la neología.

En abril de 2005, en el acto de lanzamiento público de TERMINESP, AETER manifestó que esperaba que a finales de 2006 el proyecto estaría aprobado y el proceso de trabajo en marcha. No podemos decir que se cumpliera este propósito totalmente, aunque sí de manera parcial, lo cual no es baladí si tenemos en cuenta la precariedad de recursos humanos estables con que todo el trabajo se llevó a cabo.

TERMINESP no consiguió ni reconocimiento institucional ni financiación suficiente. Y así acabó el intento.

Pero lo importante de esta historia es que felizmente veinticinco años más tarde TeresIA ha tomado el relevo de TERMINESP, con una filosofía distinta y en el contexto de las tecnologías disponibles en la actualidad. Demos pues la bienvenida a TeresIA, que se describe en detalle en esta obra, y acompañémosla para que cumpla con su cometido, porque a fin de cuentas lo que más importa es que la terminología del español y la lengua en su conjunto gocen de buena salud por muchos años.

M. TERESA CABRÉ
Barcelona, mayo de 2026

Autores

ARSENIO ANDRADES. Universidad Politécnica de Madrid. Licenciado en Traducción e Interpretación y en Filología Inglesa, y doctor en Lingüística Inglesa. En la actualidad es profesor del Departamento de Lingüística Aplicada a la Ciencia y la Tecnología de la Universidad Politécnica de Madrid, donde imparte docencia en comunicación académica y profesional a estudiantes de ingeniería. Sus líneas de investigación se centran en la traducción, la fraseología, la lingüística de corpus, el lenguaje jurídico y las lenguas especializadas. Es miembro del Grupo de Investigación UCM Traducción, Traducción Audiovisual (TAV) y Accesibilidad Lingüística (TRADAVAL), y participa en el proyecto internacional MUST de la Universidad de Lovaina. Ha trabajado en el Comité Económico y Social de la Unión Europea, y colabora puntualmente en proyectos de traducción de organismos internacionales como la UNESCO y la ONU.

CARLOS BADENES-OLMEDO. librAiry S. L. Doctor en Inteligencia Artificial, profesor del Departamento de Sistemas Informáticos de la Universidad Politécnica de Madrid (UPM) y cofundador de librAiry, *spin-off* orientada a la transferencia de tecnologías de tratamiento, organización y explotación del conocimiento. Su trayectoria se sitúa en el ámbito de la inteligencia artificial, el procesamiento del lenguaje natural, los grafos de conocimiento y los modelos de lenguaje. Ha participado en proyectos nacionales y europeos relacionados con la explotación de grandes colecciones documentales, la educación superior, la salud pública y la toma de decisiones basada en evidencia científica. Su perfil combina una sólida actividad académica e investigadora con una orientación aplicada hacia el diseño de sistemas inteligentes que facilitan el acceso, análisis y reutilización del conocimiento.

M. TERESA CABRÉ. Institut d'Estudis Catalans. Doctora en Filología Románica por la Universidad de Barcelona (1977). Ha sido catedrática de Terminología y Lingüística en la Facultad de Traducción y Ciencias del Lenguaje de la Universidad Pompeu Fabra de 1993 a 2022, y anteriormente de Lingüística descriptiva en el Departamento de Filología Catalana de la

Universidad de Barcelona (1990-1994). Es presidenta del Institut d'Estudis Catalans desde 2021. Anteriormente (2014-2021), presidió la Sección Filológica de dicha institución. Fue fundadora y primera directora del Centro de Terminología TERMCAT (1985-1989). En 1994 creó el Instituto Universitario de Lingüística Aplicada (IULA) de la Universidad Pompeu Fabra, que dirigió hasta 2004 y en cuyo marco fundó el grupo de investigación IULATERM (Léxico, terminología, discurso especializado e ingeniería lingüística) y el Observatorio de Neología, que dirigió desde su creación hasta 2014. También fue presidenta de la Red Iberoamericana de Terminología (RITERM), de 1996 a 2000; de la Asociación Española de Terminología (AETER), de 2002 a 2012, i secretaria general de la red Panlatina de Terminología (REALITER) entre 2003 i 2011. Entre otros muchos reconocimientos sociales y científicos, destaca su distinción como Chevalier de l'Ordre des Arts et des Lettres por el Ministerio de Cultura y Comunicación de la República Francesa (2008). Es doctora honoris causa por la Universidad de Ginebra, Suiza (2018) y por la Universidad de Vic (2025).

PABLO CALLEJA IBÁÑEZ. Universidad Politécnica de Madrid. Profesor permanente laboral en la Escuela Técnica Superior de Ingeniería Informática de la Universidad Politécnica de Madrid (UPM) y miembro del grupo de investigación Ontology Engineering Group (OEG). Doctorado en Inteligencia Artificial por la UPM, ha trabajado en diferentes proyectos multidisciplinares en el procesamiento del lenguaje natural aplicado a la extracción de información y representación de conocimiento en grafos. Su investigación actual se centra en el uso y mejora de los grandes modelos de lenguaje (LLM) y su integración con grafos de conocimiento y en garantizar la veracidad de las respuestas de los LLM.

MARÍA ROSA CASTRO-PRIETO. Universidad Autónoma de Madrid. Profesora del departamento de Lengua Española de la Universidad Autónoma de Madrid. Licenciada en Filología Española, Máster en Humanidades Digitales y doctora en Traducción e Interpretación, es profesora de Terminología con una trayectoria docente de más de treinta años. Su labor investigadora se centra en el léxico de especialidad y en el estudio de los procesos de formación, evolución y uso de la terminología en ámbitos científicos y técnicos. Ha trabajado como terminóloga en distintas agencias de traducción y ha colaborado en la base de datos EURODICAUTOM (actual IATE). Ha participado en diversos proyectos de investigación en áreas como la ingeniería, el medioambiente, la sanidad, el turismo y la migración. Sus publicaciones abordan tanto aspectos teóricos como aplicados de la terminología. Asimismo, cuenta con experiencia en herramientas de gestión léxica y en metodologías de lingüística de corpus.

PAULA DÍEZ-IBARBIA. Universidad Politécnica de Madrid. Lingüista computacional e investigadora, se graduó en Estudios Ingleses y obtuvo un Máster en Análisis y Procesamiento del Lenguaje, ambos por la Universidad del País Vasco. Actualmente forma parte del Ontology Engineering Group de la Universidad Politécnica de Madrid donde centra su investigación en la Web Semántica, los Datos Enlazados y el Procesamiento del Lenguaje Natural. En el marco del proyecto, sus principales responsabilidades incluyen el desarrollo de conversores para la representación de terminología en RDF (*Resource Description Framework*).

ANDREA FERNÁNDEZ VIVANCO. Universidad de Salamanca. Intérprete de conferencias e investigadora especializada en terminología y herramientas computacionales aplicadas. Es egresada del Grado en Traducción e Interpretación por la Universidad de Salamanca y del Máster en Interpretación de Conferencias por la Universidad de Granada. Su trayectoria profesional abarca el ámbito institucional (Consejo de la UE, embajadas y otros organismos internacionales), así como entornos industriales. Su investigación se centra en la terminología y las tecnologías del lenguaje, con especial interés en la aplicación de herramientas computacionales al procesamiento lingüístico. Recientemente, ha colaborado con el Proyecto TeresIA, de terminología e inteligencia artificial; la red europea ENEOLI, de neología e innovación léxica; la red NEOROC, que contrasta la neología de las distintas variedades del español; y con el proyecto E.XPER.T, que explora las vías de comunicación entre terminólogos y expertos de dominio.

MAR FORMENT FERNÁNDEZ. Universidad de Barcelona. Doctora en Lengua Española y profesora del Departamento de Filología Hispánica de la Universidad de Barcelona. Perteneció al grupo de investigación EDAP (Estudios de Discurso Académico y Profesional), que analiza la comunicación clara y los lenguajes de especialidad, con especial atención al discurso jurídico-administrativo. Participa en el proyecto «Comunicación tributaria clara. Estudio y clarificación del discurso empleado en el procedimiento de aplazamiento y fraccionamiento de deudas». Es vicedirectora de la Cátedra UB-DIBA sobre Comunicación clara aplicada a las Administraciones públicas. Ha impartido formaciones sobre comunicación clara en numerosas instituciones, como la Agencia Estatal de Administración Tributaria (AEAT), la Agencia Estatal de Seguridad Aérea (AESA), la Generalitat de Cataluña, las diputaciones de Barcelona, Castellón y Salamanca, el Ayuntamiento de Barcelona y la Consejería de Derechos Sociales del Gobierno del Principado de Asturias. Ha coordinado el volumen monográfico *Comunicación clara. Un campo emergente en investigación y transferencia en el mundo hispánico* (2024).

ADONIS GARCÍA-ANÉS. Graduada en Psicología por la Universidad Autónoma de Madrid y en Ciencia de Datos e Inteligencia Artificial por la Universidad Politécnica de Madrid. Contratada no doctora dentro del Departamento de Inteligencia Artificial en el Grupo de Ingeniería de Ontologías (OEG) de la UPM. Es desarrolladora de aplicación TermonIA para el proyecto TeresIA.

JOAQUÍN GARCÍA PALACIOS. Universidad de Salamanca / AETER. Doctor en Filología Hispánica y catedrático de universidad en el área de Traducción e Interpretación de la Universidad de Salamanca. Profesor de distintas materias relacionadas con la Lengua Española, la Lingüística, la Terminología y las lenguas de especialidad. Su investigación ha girado en torno al léxico general y especializado, desde una perspectiva diacrónica en sus comienzos, sobre la lexicografía y la lexicología después, y en los últimos treinta años sobre la terminología y la neología terminológica. La relación de todos esos ámbitos con la lengua española le ha llevado a participar en proyectos de investigación y a dirigir trabajos sobre ámbitos diversos (neurociencias, VIH/sida, tabaquismo, archivística...), a colaborar en la elaboración de varios diccionarios y a formar parte activa de redes nacionales e internacionales

centradas en la colaboración entre lenguas neolatinas y el desarrollo de su terminología. Es presidente de la Asociación Española de Terminología (AETER). Recibió en Premio «María de Maeztu» a la excelencia científica en 2026.

CECILIO GARRIGA ESCRIBANO. Universitat Autònoma de Barcelona. Catedrático de Lengua Española en la Universitat Autònoma de Barcelona. Su investigación se centra en la lexicografía, la historia de la lengua y los lenguajes de especialidad, con especial atención al léxico científico y técnico. Dirige el grupo Neolcyt (grupo de investigación en lengua de la ciencia y de la técnica), integrado en el grupo reconocido «Lexicografía y Diacronía», y ha liderado proyectos competitivos de investigación de forma continuada desde 1997. Coordina la Red Temática «Lengua y ciencia» y participa actualmente en el proyecto del Plan Nacional «Transformación digital y patrimonio lexicográfico: preservación y aprovechamiento de los datos sobre el léxico especializado (1884-1936)» para el periodo 2023-2027. Es autor de más de un centenar de publicaciones científicas y ha desempeñado diversas responsabilidades académicas y de gestión universitaria. Es miembro de AETER (Asociación Española de Terminología) y en la actualidad es presidente de AELEX (Asociación Española de Estudios Lexicográficos).

ELEA GIMÉNEZ TOLEDO. Consejo Superior de Investigaciones Científicas. Investigadora del Consejo Superior de Investigaciones Científicas (CSIC), en el Instituto de Lengua, Literatura y Antropología, y doctora en Documentación por la Universidad Carlos III de Madrid. Está especializada en evaluación de la actividad investigadora, edición académica, comunicación científica y transferencia de resultados de investigación en Humanidades y Ciencias Sociales, campos en los que ha trabajado más de veinte años. En la actualidad desarrolla investigación sobre y para el multilingüismo en la comunicación científica, siempre en el marco de la ciencia abierta y en proyectos inscritos en las humanidades digitales. Es investigadora principal del proyecto *TeresIA. Terminología e inteligencia artificial en español*, desarrollado entre cinco instituciones y que forma parte de la Estrategia Nacional de Inteligencia Artificial. Sus desarrollos están permitiendo colaboraciones muy relevantes con la ONU y con la Dirección General de Traducción de la Comisión Europea. Es consejera científica de la Fundación Lilly y presidenta del comité asesor de *The Conversation* España. Forma parte del consejo de redacción de *Revista de Occidente*.

VÍCTOR GONZÁLEZ RUIZ. Universidad de Las Palmas de Gran Canaria. Profesor de traducción jurídica en la Facultad de Traducción e Interpretación de la Universidad de Las Palmas de Gran Canaria (ULPGC), y traductor-intérprete jurado del idioma inglés desde 1995. La claridad en los textos jurídicos y administrativos, ya sean originales o traducciones, han sido el objeto principal de su labor en la enseñanza, en la investigación y en la práctica profesional. Además de participar en tareas de voluntariado como miembro de las asociaciones Clarity (de la que es su representante en España) y de la Federación Internacional de Lenguaje Claro, en la actualidad participa en proyectos dedicados al discurso tributario claro (entre 2024 y 2028, bajo la dirección de la Universidad de Barcelona) y al *legal design* (entre 2026 y 2030, como parte de una Acción COST-Asociación para la Cooperación Europea en Ciencia y Tecnología).

BEATRIZ GUERRERO GARCÍA. Universidad de Salamanca. Investigadora predoctoral gracias a un contrato de la USAL cofinanciado con el Banco Santander en la Facultad de Traducción y Documentación de la Universidad de Salamanca, mismo centro donde realizó sus estudios de grado en Traducción e Interpretación. Posteriormente, cursó el Máster en Lingüística Teórica y Aplicada en la Universitat Pompeu Fabra de Barcelona, entorno en el que se adentró en el ámbito de la lexicografía especializada y su implementación en el ámbito escolar. Actualmente, colabora con los proyectos NeoUSAL dentro de la red NEO-ROC en el estudio de neologismos detectados en prensa y TeresIA como parte del equipo lingüístico. Su investigación doctoral se centra en la dimensión conceptual de la variación terminológica en géneros textuales más allá de la comunicación especializada tradicional, como redes sociales o textos divulgativos en el ámbito de la transición energética.

MARTIN KRALLINGER. Barcelona Supercomputing Center (BSC). Dirige el equipo NLP-4BIA en el Barcelona Supercomputing Center. Es experto en minería de textos biomédicos y clínicos, así como en tecnologías del lenguaje, con más de quince años de experiencia en el desarrollo de soluciones basadas en inteligencia artificial para el ámbito de la salud. Su trabajo abarca aplicaciones en farmacovigilancia, biomateriales, cardiología, oncología, toxicología y salud laboral, entre otras áreas. Ha liderado iniciativas de evaluación y *benchmarking* de tecnologías lingüísticas biomédicas y modelos de lenguaje de gran tamaño (LLM), y ha impulsado la creación de conjuntos de datos anotados y la organización de tareas compartidas para evaluar la calidad de herramientas de detección y normalización de conceptos clínicos, extracción de información, traducción médica e indexación semántica. Es uno de los principales organizadores de campañas internacionales de evaluación de sistemas de IA y tecnologías del lenguaje, como BioCreative, y ha participado en otras iniciativas como IberLEF, Biomedical WMT, eHealth CLEF, BioASQ y BioNLP-ST. Su grupo también desarrolla corpus médicos y componentes de PLN basados en *transformers*, para lo cual utiliza la infraestructura de computación de alto rendimiento del BSC.

SALVADOR LIMA-LÓPEZ. Barcelona Supercomputing Center. Lingüista computacional e investigador en procesamiento del lenguaje natural aplicado a textos biomédicos y clínicos. Graduado en Lenguas Modernas, Cultura y Comunicación de la Universidad Autónoma de Madrid y más tarde egresado del Máster en Análisis y Procesamiento del Lenguaje de la Universidad del País Vasco, actualmente trabaja en el Barcelona Supercomputing Center. Su investigación abarca distintos ámbitos del Procesamiento del Lenguaje, con un especial foco en la extracción de información y la generación de datos anotados y protocolos de anotación. Ha colaborado como organizador de *workshops* y campañas de evaluación en iniciativas nacionales e internacionales como IberLEF, Biomedical WMT, CLEF, BioASQ o BioCreative.

NAVA MAROTO. Universidad Politécnica de Madrid / AETER. Profesora en el Departamento de Lingüística Aplicada a la Ciencia y la Tecnología de la Universidad Politécnica de Madrid, donde imparte docencia en comunicación académica y profesional a estudiantes de ingeniería. Doctora en Traducción y especialista en terminología y neología, ha dedicado su trayectoria a investigar cómo el lenguaje especializado facilita la transmisión pre-

cisa del conocimiento. Sus líneas de investigación incluyen la terminología, la neología y el uso de sistemas de aprendizaje profundo para la detección y representación del conocimiento especializado. Ha colaborado con universidades británicas en proyectos de One Health y ha realizado estancias de investigación en Canadá y Reino Unido. Además, realizó prácticas de traducción y gestión terminológica en la Dirección General de Traducción de la Comisión Europea. Actualmente es vicepresidenta de AETER, desde donde impulsa iniciativas para el fomento de la terminología, convencida del valor estratégico de los términos para la comunicación.

SERGI MARSOL TORRENT. Ingeniero biomédico y electrónico, especializado en inteligencia artificial aplicada a la salud. Trabajó durante dos años como *research engineer* en el Barcelona Supercomputing Center (BSC), desarrollando proyectos de procesamiento del lenguaje natural (NLP) y aprendizaje automático para datos biomédicos y clínicos, entre ellos TeresIA. Actualmente es becario Fulbright y cursa un Máster en *Computer Science and Artificial Intelligence* en la University of California San Diego (UC San Diego). Sus intereses se centran en la IA biomédica, los modelos fundacionales y el desarrollo de sistemas de inteligencia artificial aplicados a la investigación y la práctica clínica.

PATRICIA MARTÍN-CHOZAS. Universidad Politécnica de Madrid. Profesora ayudante doctora en la Universidad Politécnica de Madrid (UPM) y miembro del Grupo de Ingeniería Ontológica (OEG) de la misma universidad. Su investigación se ha centrado en la generación y representación de recursos terminológicos mediante tecnologías de la Web Semántica. Actualmente participa en varios proyectos de investigación donde combina estas áreas, tanto en el ámbito nacional como europeo. Sus líneas de investigación se centran en la integración y explotación de recursos terminológicos publicados en formatos estructurados de la Web Semántica (siguiendo los principios de los Datos Enlazados) para mejorar tareas de procesamiento del lenguaje natural en dominios específicos.

ELENA MONTIEL-PONSODA. Universidad Politécnica de Madrid. Profesora Titular del Departamento de Lingüística Aplicada a la Ciencia y a la Tecnología de la Universidad Politécnica de Madrid (UPM) y miembro del grupo de investigación Ontology Engineering Group del Departamento de Inteligencia Artificial de la misma universidad. Es licenciada en Traducción e Interpretación por la Universidad de Alicante, *bachelor* en Interpretación Especializada por la Hochschule Magdeburg-Stendal, Alemania, y doctora en Lingüística Aplicada por la UPM. Su investigación se centra en la conversión de terminologías, y otros recursos lingüísticos, a los formatos interoperables de la Web de Datos, para garantizar su integración en tecnologías de procesamiento de lenguaje natural. En la actualidad, coordina los trabajos de la UPM para el desarrollo de servicios y componentes avanzados de inteligencia artificial para la terminología en el proyecto TeresIA. Es investigadora principal del proyecto INESData, sobre la creación de espacios de datos, en el que se ha implementado el primer prototipo de espacio de datos lingüístico en España. Desde diciembre de 2022 representa a la UPM en el nodo CLARIAH-CM, cuyo objetivo es potenciar la investigación en Humanidades Digitales con el apoyo de las infraestructuras europeas CLARIN-ERIC y DARIAH-EU.

LUCÍA PALACIOS. Universidad Politécnica de Madrid. Investigadora predoctoral en el Ontology Engineering Group de la Universidad Politécnica de Madrid, donde desarrolla su actividad en el ámbito de la Ciencia de Datos aplicada a las Tecnologías del Lenguaje. Su investigación se centra en la intersección entre terminología, procesamiento del lenguaje natural (PLN) y Linguistic Linked Data. Su trabajo se ha especializado en el desarrollo de tecnologías semánticas orientadas a la integración de recursos lingüísticos heterogéneos, mediante técnicas de desambiguación del sentido de las palabras (WSD) y de mapeo para facilitar la transformación de terminologías especializadas a formatos estándar. Ha participado en proyectos nacionales como TeresIA, orientado a la aplicación de técnicas de inteligencia artificial a recursos terminológicos, e INESData, cuyo objetivo es el desarrollo de recursos lingüísticos y tecnologías para PLN en entornos de espacios de datos. Es graduada en Filología Hispánica por la Universidad Complutense de Madrid y posee un Máster en Análisis y Procesamiento del Lenguaje por la Universidad del País Vasco. En el marco de este máster, realizó una estancia de investigación en AHOLAB: Signal Processing Laboratory, perteneciente a la Escuela de Ingeniería de Telecomunicación de la misma universidad.

JOEL PARDO-FERRERA. Universidad Politécnica de Madrid. Investigador predoctoral en Inteligencia Artificial en la Universidad Politécnica de Madrid y miembro del Ontology Engineering Group. Es graduado en Ciencia de Datos por la Universitat de València y Máster en Inteligencia Artificial por la Universidad Politécnica de Madrid. Su investigación se centra en el Procesamiento del Lenguaje Natural, la extracción de información a partir de textos no estructurados, la extracción de relaciones y tripletas, y la construcción de recursos estructurados como grafos de conocimiento. Ha participado en proyectos como TeresIA, orientado al estudio de la terminología y la extracción de conocimiento en dominios especializados; INESData, vinculado al desarrollo de recursos lingüísticos, espacios de datos y tecnologías para el procesamiento del lenguaje; y AI4Labour, centrado en la aplicación de inteligencia artificial al análisis del mercado laboral, la automatización y el *reskilling*. Su perfil combina investigación en PLN, ingeniería de datos, recursos semánticos, transferencia tecnológica y desarrollo de soluciones aplicadas basadas en inteligencia artificial, con especial interés en la conexión entre investigación, prototipado y emprendimiento tecnológico.

MARIANO RICO. Universidad Politécnica de Madrid. Investiga en Web Semántica, Datos Enlazados, Interacción Persona-Ordenador, y Procesamiento de Lenguaje Natural. Es profesor titular de Universidad en el Departamento de Inteligencia Artificial (DIA) de la Escuela Técnica Superior de Ingenieros Informáticos (antigua Facultad de Informática) de la Universidad Politécnica de Madrid (UPM). Perteneció al Grupo de Ingeniería de Ontologías (OEG), y es el creador y responsable de la DBpedia del español. Es licenciado en Ciencias Físicas (especialista en Física Fundamental), y doctor en Ingeniería Informática y Telecomunicación, así que se atreve a meterse en cualquier campo en el que pueda ser útil. De un tiempo a esta parte trabaja con lingüistas y filólogos.

JUDITH ROSELL. Lingüista computacional especializada en el ámbito biomédico. Desde 2023, forma parte del grupo NLP for Biomedical Information Analysis (NLP4BIA) del Barcelona Supercomputing Centre (BSC), dirigido por el Dr. Martin Krallinger, donde desarrolla sistemas multilingües de Procesamiento del Lenguaje Natural aplicados a los ámbitos clínico y biomédico. Judith es graduada en Lingüística Aplicada por la Universitat Oberta de Catalunya (UOC), tiene un Máster en Análisis y Procesamiento del Lenguaje por la Universidad del País Vasco (UPV/EHU) y actualmente cursa un Máster en Ciencia de Datos en la UOC. Sus áreas de especialización incluyen la creación de conjuntos de datos anotados de alta calidad y el entrenamiento de modelos para el reconocimiento y extracción de entidades clínicas y biomédicas, su normalización para la estandarización de información médica y la clasificación de textos pertenecientes a estos dominios. Ha participado en proyectos nacionales y europeos, entre los que destacan los proyectos Horizon Europe BIOMATDB y DataTools4Heart, en los que ha colaborado con equipos multidisciplinares internacionales.

MIGUEL SÁNCHEZ IBÁÑEZ. **Universidad de Valladolid.** Licenciado en Traducción e Interpretación, Máster en Elaboración de Diccionarios y Control de Calidad del Léxico, Máster en Estudios LGTBIQ+ y doctor en Traducción y Mediación Intercultural, con mención internacional. Ha trabajado como profesor en distintas universidades españolas y estadounidenses, y como terminólogo para la Organización Mundial de la Propiedad Intelectual. Es profesor del Departamento de Lengua Española de la Universidad de Valladolid, donde coordina el Grado de Español: Lengua y Literatura. Asimismo, representa a España en el comité de coordinación de ENEOLI, la Red Europea de Estudios de Innovación Léxica. También es socio fundador de MariCorners (www.maricorners.es), asociación dedicada a la difusión de estudios y trabajos académicos de cualquier ámbito de especialidad relacionados con cuestiones LGTBIQ+.

VICTORIA SAURA-MONTESINOS. **Universidad Politécnica de Madrid.** Lingüista computacional e investigadora especializada en terminología, tecnologías del lenguaje e inteligencia artificial. Desde 2024 forma parte del Ontology Engineering Group (OEG) en la Universidad Politécnica de Madrid, donde participa en proyectos vinculados a la creación, análisis y explotación de recursos terminológicos en español, especialmente en dominios de especialidad como el jurídico-administrativo. Su trayectoria profesional se ha desarrollado en entornos nacionales e internacionales, tanto en instituciones públicas como privadas. Ha trabajado para la Unidad de Coordinación de Terminología del Parlamento Europeo, donde fue miembro del IATE Management Group. También es miembro de la Red Iberoamericana de Terminología (RITerm), red dedicada al intercambio y la cooperación en el área de la terminología, y de la COST Action ENEOLI, European Network on Lexical Innovation, orientado a metodologías y herramientas digitales aplicadas al estudio de la innovación léxica.

MARKUS SCHÜLLER PERDIGÓN. **librAiry S.L.** Graduado en Ingeniería Informática y cuenta con una trayectoria profesional enfocada en la arquitectura de sistemas y el diseño de *software* de alto nivel, tanto en entornos empresariales como en proyectos de base tecnológica

AUTORES

y de desarrollo científico. En su rol como ingeniero de *software* en librAlry, ha construido, diseñado y administrado infraestructuras escalables, orientadas al procesamiento de grandes volúmenes de datos y a la implementación de sistemas de inteligencia artificial. También ha participado activamente en el proyecto TeresIA, donde se encargó de la puesta en marcha del portal de terminologías tomando como base la tecnología de OntoPortal, adaptando y personalizando su arquitectura para cubrir los requisitos específicos del proyecto.

colección:

INTERLINGUA

427

Dirigida por:
Ana Belén Martínez López y Pedro San Ginés Aguilar

Este es un libro sobre las palabras de la ciencia y lo mucho que hay que cuidarlas para que una lengua tenga relevancia en este ámbito.

En campos como la traducción, la interpretación, la enseñanza del español o el lenguaje claro, la utilización precisa de los términos es irrenunciable. Localizar un término entre los cientos de diccionarios que existen puede no ser fácil. Disponer de terminologías actualizadas, validadas y en formatos adecuados para ser utilizadas por las tecnologías del lenguaje es crucial para que una lengua se desenvuelva bien en la ciencia y en el entorno digital. Identificar la neología de especialidad, validarla con expertos y sancionarla lingüísticamente es clave para contar con terminologías de calidad.

Este libro muestra qué investigación y qué desarrollos han sido necesarios para afrontar esos procesos críticos con los términos y para crear la infraestructura clave de la terminología en España, siempre en conexión con otras lenguas y en el marco de la ciencia abierta, contribuyendo así al multilingüismo digital.

Esta obra es para investigadores y profesionales de la lengua, pero también para todos aquellos curiosos e interesados en el trabajo que se hace con las pequeñas palabras para construir una lengua grande en ciencia.



ISBN 979-13-7033-126-9

