

Dossier: terminología en la era digital: innovación, inclusión y diversidad

El proyecto TerlexWeb en el marco del Gestor Terminológico y Documental del Instituto de Lingüística Aplicada

The TerlexWeb project within the framework of the Terminological and Documentary Manager of the Applied Linguistics Institute

Le projet TerlexWeb dans le cadre du Gestionnaire Terminologique et Documentaire de l'Institut Universitaire de Linguistique Appliquée

Sara Gómez Díaz^a ORCID: [0000-0003-4190-9256](https://orcid.org/0000-0003-4190-9256)

Alba Macías Couso^b ORCID: [0000-0001-6083-2077](https://orcid.org/0000-0001-6083-2077)

^a Universidad de Cádiz, España, Av. Dr. Gómez Ulla, 1, 11003. sara.gomez@uca.es

^b Universidad de Cádiz, España, Av. Dr. Gómez Ulla, 1, 11003. alba.macias@uca.es

Resumen

Esta contribución se enfoca en la presentación de la plataforma GETERDOC, acrónimo de Gestor Terminológico y Documental, una herramienta para el trabajo terminográfico implementada por el Instituto Universitario de Investigación en Lingüística Aplicada producto de diferentes proyectos de investigación ya completados y otro que se encuentra en fase de desarrollo. La plataforma GETERDOC está conformada por tres pilares fundamentales que apoyan algunas de las fases del trabajo terminológico: un repositorio de documentación lingüística multilingüe; una aplicación para la búsqueda de contextos dentro de un corpus metalingüístico y multilingüe, y, finalmente, TerLexWeb, un gestor de fichas terminológicas. El objetivo principal de este trabajo es realizar una descripción de la plataforma GETERDOC, su configuración y diferentes aplicaciones. Para ello, trataremos los antecedentes y la metodología desarrollada para su creación, además de mostrar sus posibilidades en el ámbito de la lingüística general y la terminología, en particular, con dos ejemplos prácticos mediante al análisis de dos términos: *lingüística del texto* y *jerga*.

Palabras clave: TERMINOLOGÍA; TERMINOGRAFÍA; LINGÜÍSTICA DIGITAL; GETERDOC; LINGÜÍSTICA DEL TEXTO; JERGA

Abstract

This contribution focuses on the presentation of the GETERDOC platform, an acronym for *Gestor Terminológico y Documental* (Terminological and Documentary Manager), a tool for terminographic work developed by the University Research Institute for Applied Linguistics (ILA). It is the result of several completed research projects and one currently in progress. The GETERDOC platform is structured around three fundamental pillars that support various phases of terminological work: a metalinguistic and multilingual repository; an application for context searching within a metalinguistic and multilingual corpus; and, finally, TerLexWeb, a manager for terminological entries. Therefore, the main objective of this work is to provide a description of the GETERDOC platform, its configuration, and its various applications. To this end, we will address the background and the methodology developed for its creation, as well as highlight its potential within the field of general linguistics and terminology, in particular, through two practical examples involving the analysis of two terms: *text linguistics* and *jargon*.

Keywords: TERMINOLOGY; TERMINOGRAPHY; DIGITAL LINGUISTICS; GETERDOC; TEXT LINGUISTICS; JARGON

Résumé

Cette contribution porte sur la présentation de la plateforme GETERDOC, acronyme de *Gestionnaire Terminologique et Documentaire*, un outil destiné au travail terminographique, mis en œuvre par l'Institut Universitaire de Recherche en Linguistique Appliquée (ILA). Elle est le fruit de plusieurs projets de recherche déjà achevés, ainsi que d'un autre en cours de développement. La plateforme GETERDOC repose sur trois piliers fondamentaux qui soutiennent différentes phases du travail terminologique : un répertoire métalinguistique et multilingue ; une application de recherche de contextes dans un corpus métalinguistique et multilingue ; et enfin, TerLexWeb, un gestionnaire de fiches terminologiques. L'objectif principal de ce travail est donc de proposer une description de la plateforme GETERDOC, de sa configuration et de ses diverses applications. À cette fin, nous aborderons les antécédents et la méthodologie développée pour sa création, ainsi que ses possibilités d'utilisation dans le domaine de la linguistique générale et, plus particulièrement, de la terminologie, à travers deux exemples pratiques portant sur l'analyse de deux termes : *linguistique textuelle* et *jargon*.

Mots-clés: TERMINOGRAPHIE, LINGUISTIQUE NUMÉRIQUE, GETERDOC, LINGUISTIQUE TEXTUELLE, JARGÓN

Fecha de recibido: 23/07/2025

Fecha de aceptado: 11/08/2025

1. La plataforma GETERDOC y el proyecto TerlexWeb

La incorporación de las nuevas tecnologías de la información a disciplinas como la Lingüística ha contribuido a la creación de nuevas herramientas y metodologías para el procesamiento y la visualización de datos. En este sentido, durante los últimos años se han desarrollado dos nuevas ramas del conocimiento: las humanidades digitales y la lingüística digital, que, junto a otras más consolidadas, como la documentación y la informática, han propiciado la creación de aplicaciones relacionadas con el tratamiento de fuentes bibliográficas y el trabajo terminográfico. Antes de comenzar con el desarrollo de la plataforma GETERDOC, es importante definir los conceptos *humanidades digitales* y *lingüística digital*, ya que son de vital importancia en este trabajo. El primero de ellos es un

nuevo campo interdisciplinario que busca entender el impacto y la relación de las tecnologías de cómputo en el quehacer de los investigadores en las Humanidades. También se conoce con los términos de Recursos Digitales para las Humanidades, Computación para las Humanidades, Computación en las Humanidades, Informática Digital y Cultural e Informática para las Humanidades (Galina Russell, 2011, p. 3).

El procesamiento de datos de carácter textual forma parte de estas relaciones entre las nuevas tecnologías y las humanidades. El concepto ha ido evolucionando a lo largo del tiempo hasta llegar a entenderse como «prácticas en torno al texto [porque] la historia, la filosofía y las letras parten de los textos y llegan a los textos; sus trabajos y sus días transcurren en la transmisión, el análisis y la producción de textos» (Rivero Franyutti, 2013, p. 98). Ante estas alusiones al texto, debemos pensar, inevitablemente, en el lenguaje como objeto de estudio, como un producto cultural y social inherente al ser humano, como un medio de transmisión del conocimiento. Por lo tanto, no es de extrañar que dentro de estas humanidades digitales haya cobrado especial importancia la lingüística como ciencia que estudia el lenguaje y las lenguas; de ahí que surja el concepto de *lingüística digital*, definido como un

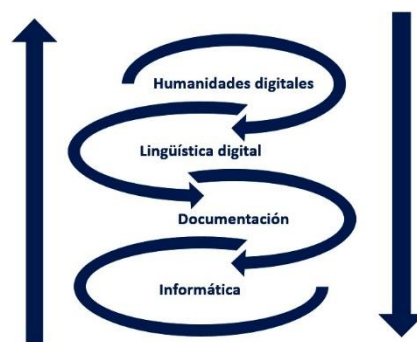
conjunto de técnicas y herramientas enfocadas a la manipulación de datos lingüísticos en formato digital, lo que incluye el almacenamiento, la representación, la manipulación y la difusión de este tipo de datos. También se ocupa de extraer y representar la información lingüística en formato

electrónico, así como las mejoras prácticas para trabajar esos datos, utilizando todos los beneficios de las tecnologías actuales (Crespo Miguel, 2020, p. 151).

Tanto en la definición de *humanidades digitales* como en la de *lingüística digital* observamos que se menciona la informática, esa «ciencia que estudia el procesamiento automático de la información» (Pareja et al., 1994, p. 17) y, por supuesto, si estamos tratando con textos, es fundamental acudir a otro ámbito de especialidad, esto es, la documentación, un «conjunto de disciplinas documentarios que estudian y ejecutan diversos aspectos del proceso documental (Archivística, Biblioteconomía, Documentación, Museología, Bibliografía, Bibliometría, etc.)» (López Yepes, 2007, p. 58).

La fusión de todas estas materias se produce de forma interdisciplinar y transdisciplinar: partimos de las humanidades digitales para desarrollar la lingüística digital, tomando como apoyo la documentación y la informática para la gestión, el procesamiento y la visualización de datos textuales y lingüísticos (imagen 1).

Imagen 1: Humanidades digitales, lingüística digital, documentación e informática



Fuente: Macías Couso (2025, en prensa)

Es en este paradigma que se desarrollan un conjunto de plataformas, proyectos y herramientas tecnológicas fundamentales para los estudios terminológicos, como son TermCat, TeresIA o GETERDOC, la plataforma que nos interesa especialmente para este artículo.

TermCat es el centro catalán de terminología. Posee multitud de funcionalidades, entre las que destacan más de 200 diccionarios en línea y los gestores

terminológicos GdT y GdTWeb, si bien no son accesibles para el público en general. No obstante, esta entidad sí ofrece Cercaterm,^[1] un buscador de fichas terminológicas multilingüe, gratuito y en línea para 32 ámbitos de conocimiento.

Por su parte, TeresIA ^[2] es un ambicioso proyecto de investigación en terminología e inteligencia artificial, con enfoque en español y las lenguas cooficiales del Estado, que generará un metabuscador de términos y otras herramientas para la extracción y gestión de neologismos. Entre sus funcionalidades, actualmente está disponible KeyCare, una interfaz implementada en español para el dominio biomédico, pero que puede ser adaptada a otros idiomas y dominios. Permite la extracción, categorización y clasificación de términos extraídos de diferentes textos.

No obstante, en esta ocasión nos centraremos en GETERDOC, herramienta cuyo nombre es acrónimo de Gestor Terminológico y Documental del Instituto de Lingüística Aplicada, pues el objetivo principal que nos marcamos en este trabajo es describir esta aplicación y sus desarrollos dentro del ámbito de la terminología lingüística, ya que, en primera instancia, se ha elaborado exclusivamente para este ámbito de conocimiento. GETERDOC es el resultado de diferentes proyectos de investigación llevados a cabo a lo largo de los años por parte del grupo de investigación Semaínein (HUM147) y el Instituto Universitario de Investigación en Lingüística Aplicada, que actualmente se encuentra inmerso en un nuevo proyecto de investigación, en el marco de la convocatoria de 2023 de Proyectos de Generación del Conocimiento. Este proyecto, denominado «Aplicaciones de la lingüística digital al ámbito de la terminología: la creación de un léxico relacional bilingüe de los usos terminológicos de la semántica léxica (TerlexWeb) (PID2022-139201OB-I00)», supone la culminación de otros tres proyectos de investigación cuya duración se extendió entre los años 2015 y 2023: 1) «Comunicación especializada y terminografía: usos terminológicos relacionados con los contenidos y tendencias actuales de la semántica léxica (TERLEX) (FFI2014-54609-P)», que se extendió durante el periodo 2015-2018; 2) «Lingüística y humanidades digitales: base de datos relacional de documentación lingüística (PY18-FR-2511)», durante los años 2020-2023, y 3) «Lingüística y nuevas tecnologías de la información: la creación de un repositorio electrónico de documentación lingüística (FEDER-UCA18-107788)»,^[3] también durante el periodo 2020-2023. Uno de los

principales hitos que derivan de la consecución de estas investigaciones se manifiesta en el registro de dos patentes intelectuales: DOCUTERM (Casas Gómez, 2020) e INTERLEX (Casas Gómez y Marín Honor, 2021). La primera corresponde a un modelo de documentación de fichas terminológicas, que contiene una serie de campos novedosos en el marco de la gestión terminológica, sobre todo el campo subíndice, en el que se trabaja a nivel de uso, el nivel de metalengua o las remisiones a relaciones conceptuales de diferentes tipos: sinonimia, antonimia o hiperonimia-hiponimia. Este modelo ha sido confeccionado especialmente para la terminología lingüística (Casas Gómez, 2022), pero sería posible aplicarlo a cualquier otro ámbito de especialidad. En cuanto a la segunda, consiste en una interfaz gráfica para la representación de relaciones terminológicas entre distintos usos.

Algunos antecedentes directos de estas investigaciones se encuentran en los trabajos de Ramírez Salado y Álvarez Torres (2018), Gómez Díaz (2022) o Álvarez Torres y Gómez Díaz (2023). Precisamente, Ramírez Salado y Álvarez Torres (2018) describían los inicios del Proyecto TERLEX, que marcó el desarrollo de esta aplicación, así como la metodología empleada para la extracción de candidatos a término. Seguidamente, Gómez Díaz (2022) mostraba los avances del repositorio lingüístico y las obras incluidas hasta la fecha en él, así como una vertiente importantísima vinculada al ámbito especializado de la semántica, pues en el repositorio se dispone de una ingente bibliografía sistematizada de este campo de especialidad, que cuenta actualmente con 21 716 referencias. Por último, Álvarez Torres y Gómez Díaz (2023) exponían los avances significativos del proyecto e incluían como ejemplo de utilidad de la plataforma y de la conformación del corpus el análisis de los usos terminológicos del concepto *semasiología*, recopilado en el modelo de ficha propuesto por Casas Gómez (2006, 2015a, 2020, 2022) e implementado en la interfaz.

Como ya hemos comentado, la plataforma GETERDOC es el resultado de todo lo descrito anteriormente, ya que combina la terminología, la informática y la gestión documental; se puede definir como

una plataforma digital destinada a la realización de trabajos de investigación lingüística, especialmente en el ámbito de la terminología, constituida por tres pilares fundamentales: una aplicación de búsqueda en corpus, un repositorio

lingüístico y una aplicación de gestión terminológica llamada TerlexWeb (Macías Couso, 2025, en prensa).

Está formada por tres pilares fundamentales, relacionados entre sí, que describiremos con más exhaustividad a continuación (imagen 2): 1) una aplicación de búsqueda en corpus, 2) un repositorio lingüístico y 3) TerlexWeb.

Imagen 2: Esquema de GETERDOC



Fuente: Macías Couso (2025, en prensa)

1.1. Búsqueda en corpus

La búsqueda de contextos terminológicos definitorios es una de las fases más importantes a la hora de localizar y delimitar usos conceptuales de unidades terminológicas. Esta aplicación de búsqueda en corpus está conformada por 104 673 634 unidades léxicas pertenecientes a más de 14 000 documentos en texto plano. La herramienta funciona como un buscador en el que el investigador puede localizar contextos sobre las unidades terminológicas que desee, aplicando diferentes parámetros (imagen 3):

- a) unidades terminológicas simples o compuestas,
- b) concurrencia de varios términos en un mismo contexto,
- c) exclusión de unidades terminológicas y, por último,
- d) inclusión de otras unidades terminológicas.

En la imagen 3, se muestra una captura de pantalla de las funciones descritas.

Imagen 3: Aplicación de búsqueda de contextos en el corpus en texto plano

The screenshot shows a web interface titled 'Opciones de búsqueda en Corpus'. It features three main sections for defining search criteria:

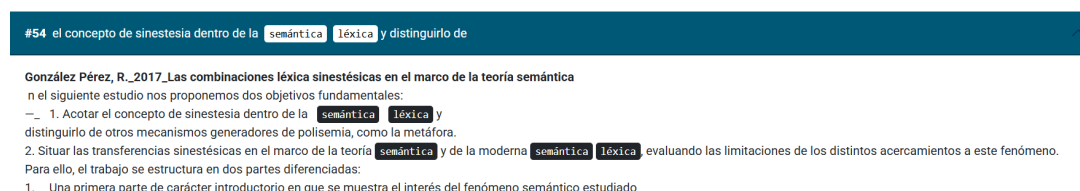
- + Añadir (Green):** Labeled 'Texto que debe aparecer' (Text that must appear). It contains a text input field with the placeholder 'Añadir términos...'. A green border highlights this section.
- Eliminar (Red):** Labeled 'Texto que no debe aparecer' (Text that must not appear). It contains a text input field with the placeholder 'Añadir términos...'. A red border highlights this section.
- ▼ Opcional (Blue):** Labeled 'Texto que puede aparecer' (Text that may appear). It contains a text input field with the placeholder 'Añadir términos...'. A blue border highlights this section.

At the top right, there is a small instruction: 'Presione **Enter** en un término para añadir varios'.

Fuente: elaboración propia

Todas estas opciones para realizar la búsqueda permiten que la herramienta se adapte al investigador en función de sus necesidades. Una vez introducido el término o unidades terminológicas, la aplicación devuelve todos los contextos en los que aparezca, ordenados por índices de aparición, en una o varias páginas, según la cantidad de ocurrencias (imagen 4). Además, en cada contexto se incluirá la referencia bibliográfica de la que se ha extraído.

Imagen 4: Ejemplo de resultados de búsqueda



Fuente: elaboración propia

Finalmente, también se le permite al usuario descargar los resultados de la búsqueda en formato de texto plano para que puedan ser visualizados o manipulados, si se considera necesario, con editores de texto.

Esta aplicación es realmente útil para el trabajo terminológico-terminográfico en el campo lingüístico, ya que en este tipo de investigaciones actuales es fundamental apoyarnos en la lingüística de corpus porque, como afirman Díaz Negrillo y Fernández Domínguez (2010, p. 38):

En este terreno los corpus son herramientas especialmente útiles puesto que no constituyen únicamente un conjunto de textos con posibles candidatos a términos; por el diseño de los mismos, proporcionan además otro tipo de

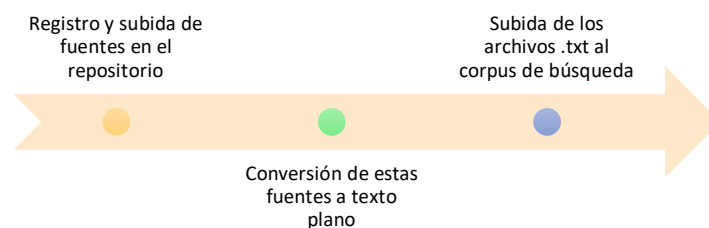
información (contexto, información lingüística, información sobre fuentes, etc.) y posibilidades metodológicas (listas, frecuencias, combinaciones de palabras, etc.), que permiten una actividad terminológica más sistemática, fiable y amplia.

Concretamente, en el ámbito terminológico-terminográfico recurrimos, entre otras funciones, a la lingüística de corpus para extraer de forma semiautomática los términos de una disciplina lingüística o de cualquier otro campo del saber científico tras la confección de un corpus, o también para la búsqueda de contextos que avalen los usos terminológicos y definiciones de los términos sometidos a análisis, que es, precisamente, lo que nos permite esta aplicación. De hecho, «el uso de corpus textuales de especialidad ha dado a la terminología la posibilidad de observar los términos como unidades del discurso» (Cabré y Da Cuhna, 2022, p. 192), en los que el contexto es fundamental para determinar la terminologicidad de una unidad léxica.

1.2. Repositorio lingüístico

El manejo de fuentes bibliográficas es otro de los pilares fundamentales dentro de la gestión terminográfica. En este sentido, cobra especial relevancia la creación de esta base de datos bibliográfica, que contiene documentación metalingüística y multilingüe. Como su propio nombre indica, está conformada por fuentes lingüísticas y filológicas (diccionarios de lingüística, artículos científicos, revistas científicas, capítulos de libros, libros y reseñas) en diferentes idiomas: alemán, árabe, aranés, asturiano, catalán, checo, esloveno, español, euskera, finlandés, francés, gallego, griego, inglés, italiano, japonés, ladino, latín, portugués, romaní, rumano y ruso. Estas fuentes son exactamente las mismas que conforman el corpus en texto plano del que hablamos en el apartado anterior, por lo que ambas aplicaciones están interrelacionadas de la siguiente manera (imagen 5):

Imagen 5: Interrelación entre el corpus y el repositorio lingüístico



Fuente: elaboración propia

El repositorio permite visualizar y descargar, en caso de que no existan derechos de autor, los documentos depositados (imagen 6). Por lo tanto, al obtener la referencia del documento en la búsqueda de corpus, el usuario puede dirigirse con estos datos al repositorio y localizar el documento original digitalizado,^[4] tal y como puede verse en el siguiente ejemplo, en el que hemos utilizado la referencia perteneciente al contexto que mostramos anteriormente:

Imagen 6: Búsqueda en el repositorio



Fuente: elaboración propia

En la imagen 6 se observa, además, que existe la posibilidad de realizar búsquedas utilizando diferentes filtros. Esto se debe a que, a la hora de registrar y volcar los documentos a la plataforma, se cumplimentan, a su vez, diferentes metadatos: materias lingüísticas en las que se enmarca el documento, palabras clave, autor(es), idioma(s) y, por último, lugar y editorial, si procede.

La actualización del repositorio y, por lo tanto, la del corpus, se realiza diariamente, de forma progresiva y simultánea. De este modo, ambas aplicaciones se encuentran

siempre actualizadas con nueva documentación; en la actualidad, cuenta con más de 14 000 documentos registrados. La actualización del corpus la realiza de forma manual el personal del Instituto Universitario de Investigación en Lingüística Aplicada, pasando por las distintas fases requeridas (véase nota 2).

1.3. TerlexWeb

El último pilar que sustenta a la plataforma GETERDOC es TerlexWeb, una aplicación destinada a la consulta, creación y gestión de fichas terminológicas. Inicialmente, se presentó como la interfaz gráfica INTERLEX (Casas Gómez y Marín Honor, 2021), pero, tras varias transformaciones, se ha optado por una nueva nomenclatura. La base metodológica de TerlexWeb se encuentra en el modelo de documentación de usos terminológicos DOCUTERM (Casas Gómez, 2022) que, como ya hemos comentado, dio como resultado una patente intelectual registrada, al igual que INTERLEX.

Un antecedente de la creación de este modelo ya aparece en Casas Gómez (2006), trabajo en el que se realiza una revisión crítica de los modelos terminográficos existentes hasta la fecha. Con el paso de los años, las críticas realizadas por el autor han ido tomando forma en la siguiente premisa:

Resulta, por tanto, imprescindible la creación y diseño de nuevo modelo dinámico, basado en el uso comunicativo y en las relaciones conceptuales de las unidades terminológicas, y no en su propia naturaleza estática como términos, que deshaga la ambigüedad en la comunicación científica y, para ello, consigne los diferentes usos y subusos de las correspondientes unidades científicas, esto es, sirva de actualización de las distintas acepciones terminológicas de todos los términos y conceptos relativos a cualquier campo de conocimiento científico, con aplicación específica al ámbito de la lingüística (Casas Gómez, 2022, p. 13).

Aunque puede ser adaptado a terminologías pertenecientes a otras áreas del conocimiento, este modelo ha sido concebido fundamentalmente para terminología lingüística. En Casas Gómez (2022, pp. 14-16), se describen los campos que conforman la ficha terminológica, que enumeramos a continuación:

- 1) término de entrada,

- 2) subíndice,
- 3) categoría lingüística del término,
- 4) fuente del término,
- 5) estatus del término,
- 6) variación denominativa,
- 7) área(s) temática(s),
- 8) etimología,
- 9) mecanismo de formación del término,
- 10) definición,
- 11) fuentes doctrinales de la definición,
- 12) nivel de metalengua del uso terminológico,
- 13) contexto representativo de cada uso terminológico,
- 14) fuente de cada contexto,
- 15) ejemplos en diferentes lenguas,
- 16) remisiones a usos sinónimos o parásinónimos de carácter terminológico,
- 17) remisiones a usos hipernímicos, hiponímicos o meronímicos de carácter terminológico,
- 18) remisiones a usos antonímicos de carácter terminológico,
- 19) coincidencias terminológicas,

20) equivalencias en otras lenguas,

21) fuentes de las equivalencias,

22) fuentes documentales del término,

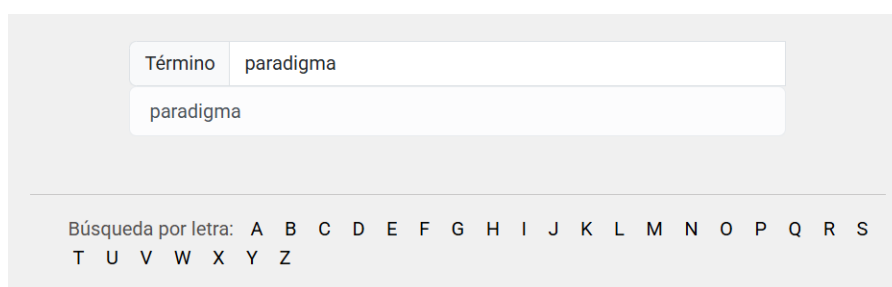
23) bibliografía recomendada,

24) observaciones y notas e

25) información y gestión de la base de datos.

La confección y representación de las fichas terminológicas de TerlexWeb siguen las pautas de este modelo terminográfico. Los diferentes ficheros pueden consultarse utilizando un buscador, al igual que en el corpus textual, introduciendo la unidad terminológica que se desee o bien buscando por la inicial, con lo que se desplegará una lista de términos ordenados alfabéticamente (imagen 7).

Imagen 7: Búsqueda de términos en TerlexWeb



Término paradigma

paradigma

Búsqueda por letra: A B C D E F G H I J K L M N O P Q R S
T U V W X Y Z

Fuente: elaboración propia

Al consultar un término, la aplicación devuelve una lista de los distintos usos o subíndices de esa unidad junto a su definición (imagen 8); se puede, además, utilizar diferentes filtros para la localización de usos terminológicos: fuente del término, estatus, nivel de metalengua y área temática.

Imagen 8: Listado de usos terminológicos para el término paradigma

Por Fuente de Término
 Todas

Por Estatus
 Todas

Por Nivel de Metalengua
 Todas

Por Áreas Temáticas
 Todas

Buscar
 Quitar Filtros

- paradigma1**
 cada uno de los modelos, incluídas todas sus formas, de la flexión nominal y verbal
- paradigma2**
 conjunto de unidades que se pueden conmutar entre sí, es decir, que pueden manifestarse en un mismo contexto
- paradigma3**
 conjunto de realizaciones científicas universalmente reconocidas, que, durante cierto tiempo, proporcionan modelos de problemas y soluciones, métodos, teorías y metas investigadoras a una comunidad científica
- paradigma4**
 respuestas asociativas que los hablantes hacen cuando reciben una palabra estímulo

Fuente: elaboración propia

Finalmente, acudiendo a cada uno de los usos, podemos ver la ficha terminológica completa (imagen 9).

Imagen 9: Ficha terminológica de paradigma

Ficha terminológica de paradigma

Información básica
 Definición
 Contextos
 Relaciones conceptuales
 Equivalencias
 Fuentes documentales y doctrinales
 Otros datos

Término	paradigma
Categoría lingüística	sustantivo
Fuentes	corpus de textos especializados fuentes lexicográficas
Estatus	documentado lexicográficamente
Áreas temáticas	semántica morfología sintaxis lingüística general fonética fonología

Fuente: elaboración propia

En definitiva, podemos afirmar que la plataforma GETERDOC representa, en conjunto, una estación de trabajo completa para cualquier terminólogo que desarrolle sus investigaciones dentro de la lingüística, ya que se sustenta sobre tres fases fundamentales de la técnica terminográfica: la localización de contextos de

uso, la gestión de fuentes bibliográficas y, por supuesto, la confección de fichas terminológicas.

El proyecto TerlexWeb se enmarca en el paradigma actual, en el que la terminología se ha convertido en un campo de estudio tanto teórico como práctico, que goza de una gran vigencia dentro de las múltiples aplicaciones de la lingüística. Este interés se debe, principalmente, a tres motivos: 1) la vinculación de la terminología a la política lingüística de aquellos países interesados en potenciar los lenguajes de distintas áreas especializadas; 2) su importancia dentro del desarrollo de aplicaciones y herramientas en el ámbito de las tecnologías del lenguaje y las industrias de la lengua, y 3) las posibilidades de transferencia empresarial y social, con el impacto que esto conlleva. Guerreros Ramos y Pérez Lagos (2021) ya destacaron el impulso de las investigaciones terminológicas dentro del panorama universitario y académico que se llevaban a cabo en España y, concretamente, en Andalucía. Desde la Universidad de Cádiz, el devenir del grupo de investigación Semaínein (HUM147) y del Instituto Universitario de Investigación en Lingüística Aplicada ha encontrado en la terminología uno de los ejes más importantes en investigación aplicada y transferencia del conocimiento (Casas Gómez, 2015b, 2021).

Como ya hemos descrito en el apartado anterior, la plataforma GETERDOC se ha creado gracias a la consecución de distintos proyectos de investigación, antesala del proyecto TerlexWeb, en el que se enmarca este trabajo. Este nuevo proyecto se centra en el marco de la semántica léxica, y adopta postulados teóricos y marcos metodológicos para el estudio descriptivo de la terminología (L'Homme, 2020), bajo la pretensión de delimitar y visibilizar los distintos usos terminológicos de la semántica. De esta manera, con TerlexWeb se persiguen dos objetivos principales:

- 1) crear un léxico relacional bilingüe español-inglés con equivalencias en otras lenguas, en formato digital, de los usos terminológicos vinculados a los nuevos contenidos y perspectivas actuales de la semántica léxica y
- 2) confeccionar una plataforma digital que funcione como herramienta de consulta, trabajo e investigación para lingüistas, traductores y correctores

lingüísticos, así como a estudiantes, investigadores y filólogos de distintas lenguas, en la que se vincule todo el material disponible en el repositorio electrónico con los distintos usos terminológicos consignados en las fichas terminológicas: a través de la búsqueda de un término, se obtendrán sus usos correspondientes y, además, se podrá acceder al corpus y consultar el material bibliográfico donde se registren dichos usos.

Estos dos objetivos dejan patente la vinculación entre GETERDOC y el proyecto TerlexWeb, ya que, sin la existencia de esta plataforma, no existiría ninguna base sólida sobre la que construir los pasos necesarios para lograr nuestro propósito y, a su vez, el trabajo de investigación terminológica que supondría la consecución del primero de los objetivos contribuye al desarrollo de la implementación de mejoras en la plataforma digital.

2. Aplicaciones terminológicas

A continuación, presentamos dos trabajos de investigación realizados dentro del marco del proyecto TerlexWeb, para ejemplificar la utilidad de la plataforma GETERDOC y sus aplicaciones dentro del ámbito de la terminología. En primer lugar, en el campo de la lingüística textual y discursiva, se propone el análisis del concepto *lingüística del texto*, acuñado por Coseriu (1955-1956), por presentar varios problemas terminológico-conceptuales en esa área de especialidad. En segundo lugar, se realiza un análisis terminológico de los conceptos variacionistas de *argot* y *jerga*, dos unidades terminológicas que, tradicionalmente, han sido utilizadas de forma sinonímica sin realmente serlo, ya que las relaciones entre ambos términos y sus subusos guardan una gran complejidad.

2.1. Usos terminológicos del término *lingüística del texto*

La unidad terminológica *lingüística del texto* ha recibido multitud de denominaciones desde su surgimiento: *lingüística superoracional*, *lingüística supraoracional*, *lingüística transfrástica*, *hipersintaxis*, *gramática textual* y, más recientemente, incluso se la ha identificado con el análisis del discurso.

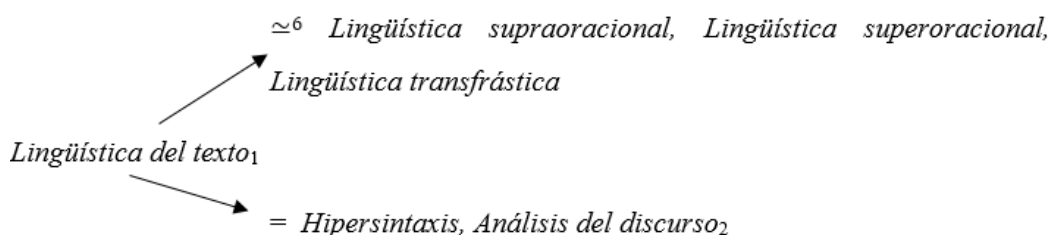
Precisamente, por este motivo, surge la necesidad de establecer límites precisos entre estas disciplinas: de ahí que recurramos a la terminología y a la lingüística de corpus para esclarecer esta problemática.

En este sentido, para analizar este concepto en el marco de su ámbito de especialidad, se confeccionó un subcorpus sobre lingüística textual y discursiva con las obras incluidas de este campo en el corpus de la plataforma GETERDOC. De esta manera, se seleccionaron 100 obras representativas en español y se consultó en ellas el concepto que ahora nos ocupa: *lingüística del texto* y su variante formal *lingüística textual*. Como resultado, obtenemos 1369 contextos para la primera forma y 411 para la segunda. Una vez seleccionadas los contextos definitorios, identificamos hasta siete usos terminológicos diferentes, algunos de ellos ya comentados en Gómez Díaz (2024a):

1. *Lingüística del texto*₁: «Es una disciplina nacida a partir de la ampliación de las gramáticas oracionales (especialmente la GGT), ello con la intención de describir y, en última instancia, de integrar en el sistema lingüístico una unidad perteneciente a un nivel superior de análisis, el *texto*» (Fernández Smith, 2007, p. 18).
2. *Lingüística del texto*₂: «Disciplina lingüística que tiene como tarea fundamental la descripción y constitución del texto» (Lewandoski, 1982, p. 215).
3. *Lingüística del texto*₃: En este tercer uso, aportamos nuestra definición propia del concepto: «La lingüística del texto es la disciplina que se ocupa del estudio de los textos como unidades comunicativas completas, incluyendo su interpretación, descripción y análisis, así como el estudio de las tipologías y clases textuales» (Gómez Díaz, 2024b, p. 395).
4. *Lingüística del texto*₄: «Se ocupa del ámbito lingüístico constituido por los actos de habla, o el entramado de los actos de habla, que realiza un determinado hablante en una situación determinada, y que puede estar integrado por manifestaciones habladas o escritas» (Casado Velarde, 2000, p. 9).

5. *Lingüística del texto*₅: «La lingüística del texto, por su parte, obedece, como sabemos, a una larga tradición de estudios orientados hacia un tipo de texto, el texto literario» (Fernández Smith, 2003, p. 235).
6. *Lingüística del texto*₆: «La Lingüística del texto, según Adam (2005, p. 19), es una disciplina auxiliar del análisis del discurso, pero, mientras la primera se ocupa preferentemente de las condiciones de subjetividad y propiedades de los textos, el análisis del discurso se asocia habitualmente a las condiciones de producción en relación» (López Alonso, 2014, p. 220).
7. *Lingüística del texto*₇: Existe un consenso generalizado respecto a que la lingüística textual estudia los textos escritos (Loureda y Schrott, 2021), por lo que puede concebirse esta disciplina como «una rama de la lingüística interesada por la tipología textual y la confección de textos, principalmente los escritos» (Gómez Díaz, 2024b, p. 403).

En cuanto a los usos destacados,[\[5\]](#) el primero de ellos concibe la lingüística del texto como la disciplina lingüística que describe el sistema lingüístico superior a la oración o la frase (el texto). Es, quizás, el uso más extendido, estandarizado y problemático, ya que presenta numerosas relaciones conceptuales y coincidencias terminológicas. Así pues, este primer uso es sinónimo de *lingüística superoracional*, *lingüística supraoracional* y *lingüística transfrástica*. A su vez, es una coincidencia terminológica de *hipersintaxis* y *análisis del discurso*₂. Apuntamos, además, el término *hipersintaxis* como una coincidencia terminológica de este uso porque es un concepto propuesto por los miembros de la Escuela de Praga —Palek (1968) fue el precursor— para referirse a la disciplina de la lingüística del texto. Asimismo, señalamos también como coincidencia terminológica *análisis del discurso*₂, ya que, en los inicios de esta disciplina, Harris (1952) acuña el término con este sentido:[\[6\]](#)

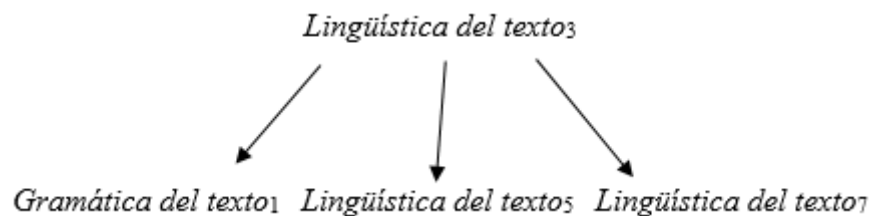


Los dos siguientes usos del término son muy próximos al concepto *gramática del texto*. De hecho, cada uno de ellos establece una relación conceptual con este término. Así pues, el segundo uso, de corte gramatical, identifica a la lingüística

del texto con la disciplina lingüística que se encarga de la descripción y conformación del texto; de ahí que se considere sinónimo de *gramática del texto*₂, que tiene la misma función:

$$\text{Lingüística del texto}_2 \simeq \text{Gramática del texto}_2$$

En el tercer uso, incluimos nuestra propia definición del concepto, con la que marcamos el límite con otras disciplinas próximas. De este modo, concebimos la lingüística del texto como «la disciplina lingüística que se ocupa del estudio de los textos como unidades comunicativas completas, incluyendo su interpretación, descripción y análisis, así como el estudio de las tipologías y clases textuales» (, 2024b, p. 80). En este sentido, basándonos en esta definición, la lingüística del texto debe encargarse del estudio de todos los textos, por lo que señalamos como hipónimo de este uso el de *lingüística del texto*₅ y el de *lingüística del texto*₇, ya que, según la tradición, la lingüística del texto estaba orientada al estudio de dos tipos de textos concretos: el literario y el escrito, que hacen alusión, respectivamente, a los usos mencionados. Asimismo, este uso también es hiperónimo de *gramática del texto*₁, pues en el sentido estricto del término, se interesa por los procesos lingüísticos de formación y constitución de los textos:



Para la cuarta acepción, destacamos un uso del concepto de corte comunicativo y pragmático, cuyo objeto de estudio es el acto de habla. Tanto la lingüística del texto como la pragmática se han preocupado por la confección y producción de textos en actos comunicativos; de ahí su estrecha relación, hasta el punto de que surge la denominación *pragmática textual*, con la que mantiene relación conceptual de sinonimia:

$$\text{Lingüística del texto}_4 \simeq \text{Pragmática del texto}$$

Por último, abordamos el sexto uso, que en este caso se aproxima de nuevo al concepto de *análisis del discurso*. Ahora se concibe la lingüística del texto como una rama del análisis del discurso encargada del estudio de las propiedades de los textos, por lo que es su hipónimo:

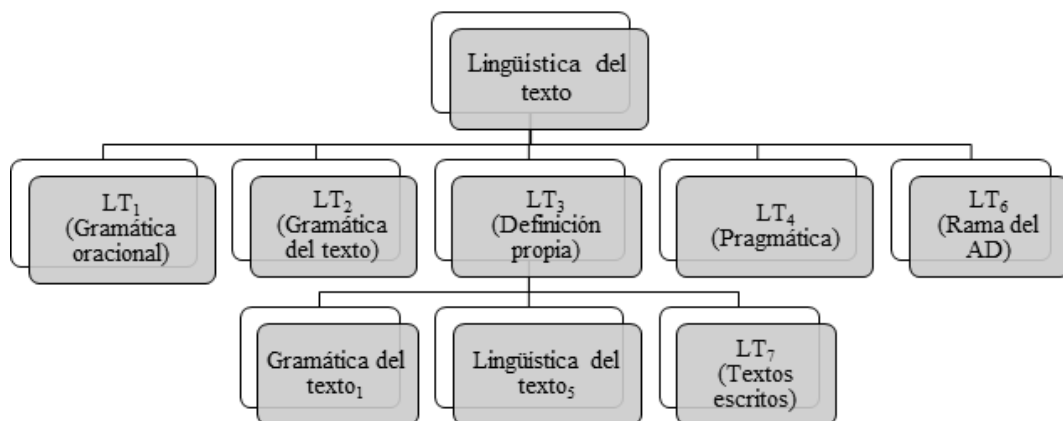
Análisis del discurso₅



Lingüística del texto₆

Para ejemplificar las redes conceptuales establecidas entre los usos de este uso terminológico, incluimos el siguiente gráfico (imagen 10):

Imagen 10: Relaciones conceptuales de la unidad terminológica lingüística del texto



Fuente: elaboración propia

2.2. Relaciones conceptuales entre los conceptos variacionistas de *argot* y *jerga*

Metodológicamente, hemos realizado las tareas requeridas por cualquier trabajo de corte terminográfico. El primer paso es acudir a las fuentes doctrinales — diccionarios especializados en lingüística— para consultar los distintos usos documentados en estas obras. El segundo paso es realizar la búsqueda en el corpus metalingüístico para localizar contextos definitorios del término y, además, descubrir otros usos posibles que no estén documentados en las fuentes doctrinales del término.

Así, para *jerga* se localizaron tres usos diferentes en las fuentes doctrinales consultadas:

1. *Jerga*₁: «Lenguajes formados en un círculo lingüístico especial» (Abraham, 1981, p. 265).
2. *Jerga*₂: «Lengua de mal gusto, compleja o incomprensible» (Alcaraz Varó y Martínez Linares, 1997, p. 312).
3. *Jerga*₃: «Conjunto de palabras de etimología oscura o difícil de hallar, que se introduce en la conversación de todos los grupos sociales por broma o ironía» (Alcaraz Varó y Martínez Linares, 1997, p. 312).

Tras realizar la consulta en el corpus de GETERDOC, la plataforma devolvió un total de 457 contextos, que se descargaron y procesaron con el *software* informático Sublime Text para corregir posibles erratas derivadas de la conversión de los documentos originales digitalizados a OCR. Tras un primer cribado de estos contextos, se seleccionaron aquellos que fuesen definitorios, incluyendo en el modelo de ficha tres contextos representativos para cada uso.

En cuanto a *argot*, en las fuentes doctrinales hemos localizado tres usos distintos:

1. *Argot*₁: «Lenguaje especial de los marginados, provisto de numerosas metáforas y deformado fonética y formalmente» (Abraham, 1981, p. 68).

2. *Argot*₂: «Subsistema léxico no estándar, de carácter eminentemente coloquial, formado principalmente por sinónimos de palabras y expresiones de la lengua estándar. Aunque no siempre aparece ligado a un dialecto y es distinto del léxico científico-técnico de las jergas, el argot muestra cierta variación regional y, con frecuencia, surge en círculos profesionales, pues forma parte de la estilística inherente a toda lengua o dialecto. Así, por ejemplo, podemos hablar de un argot barcelonés o madrileño, del argot de los estudiantes, de los músicos de rock o del mundo de la prostitución, pero también de un argot general o común que ha pasado al dominio de la lengua y es empleado por la mayoría de los hablantes en contextos coloquiales» (Alcaraz Varó y Martínez Linares, 1997, p. 64).

3. *Argot*₃: «Galicismo por el que se alude a *jerga* o *germanía*» (Cerdà Massó, 1986, p. 26).

La plataforma devolvió un total de 627 contextos para este término, que recibieron el mismo tratamiento que los de *jerga*. En este caso, encontramos que existe un uso de *argot* que se corresponde con un fenómeno variacionista en el que intervienen tanto elementos diastráticos como diatópicos, cuya característica principal es la «marca social» que lo distingue del resto de *jergas* o *lenguajes sectoriales*. Rodríguez Díez (1981, p. 54) acuña el término *signum social* como rasgo propio de los *argots*, en un esquema en el que propone diferentes criterios para diferenciar entre *argot*, *jerga* y *lenguaje científico-técnico*. Además, propone la siguiente clasificación (Rodríguez Díez, 1981, p. 53):

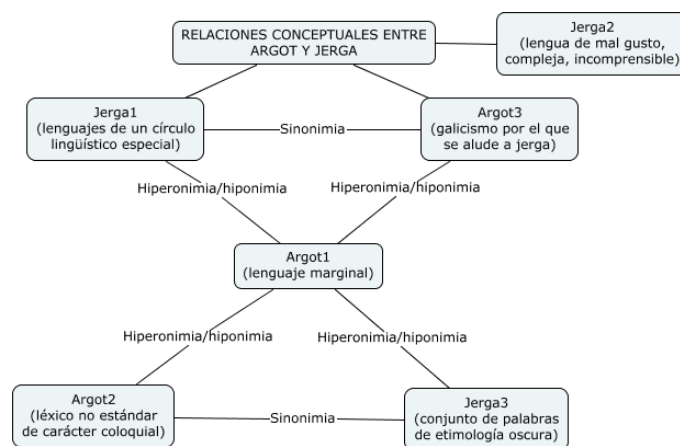
- a. El *argot* de los grupos sociales cuyas actividades se realizan al margen de la sociedad. Nos referimos, de una forma bastante restringida, a lenguas esotéricas, que tienen finalidad críptica, es decir, pretenden aislar a los no iniciados en ella.
- b. Los *lenguajes sectoriales*, las distintas *jergas* pertenecientes a profesiones u oficios afines. En este caso, no existe finalidad críptica, ya que la intención no es aislar a los demás de su contenido. De hecho, es común que los vocabularios pertenecientes a ciertas jergas se trasladen al lenguaje coloquial gracias a su difusión, como ocurre, por ejemplo, con la prensa deportiva.

c. Los *lenguajes científico-técnicos*, que abarcan las nomenclaturas pertenecientes a disciplinas científicas. En este caso, ciertas terminologías pueden resultar incomprensibles para los no iniciados, pero, en este caso, esto responde al propio carácter científico donde se enmarcan, no a la necesidad de encriptar el lenguaje.

Como observamos, el principal problema que presentan *argot* y *jerga* es su carácter polisémico, sumado a haber sido utilizados como sinónimos sin realmente serlo. Anteriormente, hemos indicado que las relaciones conceptuales entre ambos términos son más complejas y abarcadoras que una relación sinonímica y, si atendemos a las definiciones y a los contextos de uso, podemos establecer las siguientes conclusiones (imagen 11):

1. Entre *jerga₁* y *argot₃* existe una relación conceptual de sinonimia y ambos funcionan como hiperónimos generales del resto de usos terminológicos.
2. *Argot₁* es hipónimo de *jerga₁* y de *argot₃* y, a su vez, es hipéronimo de *argot₂* y de *jerga₃*.
3. Los usos de *argot₂* y *jerga₃* tienen una relación conceptual sinonímica y son, a su vez, hipónimos de *argot₁*.

Imagen 11: Relaciones conceptuales entre *argot* y *jerga*



Fuente: elaboración propia

3. Conclusiones

El avance de las nuevas tecnologías ha desempeñado un papel fundamental en el desarrollo y fortalecimiento de los estudios terminológicos-terminográficos actuales, al integrarse de forma sinérgica con otras disciplinas hoy día consideradas transversales, como son la documentación y la informática, al seguir enfoques metodológicos novedosos basados en las humanidades digitales, pero especialmente en la lingüística de corpus y la lingüística digital. Esta convergencia inter y transdisciplinar ha facilitado el acceso, la gestión y el análisis de grandes volúmenes de datos lingüísticos que permiten recopilar, almacenar y difundir terminología de distintos campos de conocimiento. En este sentido, GETERDOC supone una plataforma única que sirve de apoyo al trabajo terminográfico para la terminología lingüística mediante la unificación de sus tres pilares fundamentales: la búsqueda de contextos en un corpus metalingüístico y multilingüe, la gestión bibliográfica en un repositorio de las mismas características y, por último, la elaboración de fichas siguiendo el modelo DOCUTERM y su materialización en la interfaz gráfica INTERLEX, posteriormente denominada TerlexWeb, eje central del proyecto sobre el que hablamos en esta publicación.

En la actualidad, existen otras herramientas digitales de aplicación terminológica, como las ya comentadas (Termcat o TeresIA), que suponen, de igual modo, recursos imprescindibles para cualquier terminólogo y abarcan, además, las ontologías de distintas disciplinas. Sin embargo, la plataforma GETERDOC, además de contener un repositorio multilingüe en 23 idiomas,[\[7\]](#) lo cual permite y facilita la búsqueda de equivalentes traductológicos, presenta la particularidad de que aglutina tres aplicaciones que son imprescindibles para la realización de cualquier trabajo terminológico en el ámbito lingüístico, por lo que se trata de una herramienta fundamental para el estudio de este ámbito de conocimiento. Además, en GETERDOC, toda la información está relacionada, esto es: los textos que conforman el corpus de búsqueda se encuentran igualmente registrados en el repositorio, lo que significa que, al introducir un término y revisar sus contextos, podemos acudir directamente al repositorio para consultar la referencia

bibliográfica original de la que se extrajo ese contexto. Asimismo, de la selección de contextos definatorios se extraerán las informaciones que se recogen en la ficha terminológica, cuya bibliografía también se encuentra depositada en la base de datos. GETERDOC no puede definirse como un diccionario o una aplicación online para realizar consultas sobre terminología lingüística: se trata de la materialización de una metodología de trabajo en una aplicación que unifica documentación, terminología, lingüística de corpus e informática. Por otra parte, la adaptación de este modelo a otros ámbitos de especialidad es plausible. La aplicación se ha desarrollado en el lenguaje de programación Python, ampliamente utilizado en procesamiento del lenguaje. Con la migración de la interfaz gráfica y la subida de bibliografía especializada, podría utilizarse prácticamente en cualquier área del conocimiento.

En definitiva, gracias a estas herramientas tecnológicas, es posible realizar estudios terminológicos más dinámicos, precisos y contextualizados. Por tanto, recursos como estos se convierten en una herramienta de trabajo imprescindible para la realización de investigaciones dentro de nuestro ámbito, especialmente cuando nos centramos en la terminología lingüística. Muestra de ello son las elaboraciones de fichas terminológicas incluidas sobre los ámbitos especializados de la lingüística textual y discursiva y la variación léxica, en las que se identifican diferentes usos terminológicos de los conceptos *lingüística del texto*, *jerga* y *argot*, así como sus relaciones conceptuales y otro tipo de información fundamental para la lingüística en general y los sectores especializados comentados en particular.

Referencias bibliográficas

- Abraham, W. (1981). Diccionario de terminología lingüística actual. Gredos.
- Alcaraz Varó, E. y Martínez Linares, M. A. (1997). Diccionario de lingüística moderna. Ariel.
- Álvarez Torres, V. y Gómez Díaz, S. (2023). Terminología y documentación: la confección de un corpus metalingüístico y multilingüe de semántica léxica. En M. Casas Gómez y C. Varo Var (Coords.), *La terminología de la semántica* (pp. 77-93). Universidad de Sevilla.

- Cabré, M. T. y Da Cuhna, I. (2022). El papel de los corpus en la terminología: una mirada específica a la terminología del español. En G. Parodi, P. Cantos-Gómez y C. Howe (Eds.), *Lingüística de corpus en español / The Routledge Handbook of Spanish Corpus Linguistics* (pp. 190-204). Routledge.
- Casado Velarde, M. (2000). *Introducción a la gramática del texto del español*. Arco.
- Casas Gómez, M. (2006). Modelos representativos de documentación terminográfica y su aplicación a la terminología lingüística. *Revista de Lingüística y Lenguas Aplicadas*, 1, 25-36.
- Casas Gómez, M. (2015a). Semántica léxica, terminología y terminografía: a propósito de los usos terminológicos relacionados con los contenidos y perspectivas actuales de la semántica léxica. En J. Santana Marrero (Dir.), M. León-Castro Gómez y A. Zerva (Eds.), *Sociolingüística andaluza*, 17: *La variación en el español actual. Estudios dedicados al profesor Pedro Carbonero* (pp. 91-110). Universidad de Sevilla.
- Casas Gómez, M. (2015b). Lingüística e investigación terminológica en Andalucía. *Terminàlia*, 12, 18-28.
- Casas Gómez, M. (2020). DOCUTERM. Modelo de documentación de usos terminológicos. Patente registrada con número 202099908175522, expediente CA-0256-2020 y número de asiento registral 04/2020/4177.
- Casas Gómez, M. (2022). DOCUTERM: modelo de documentación de usos terminológicos. *Revista de Lingüística y Lenguas Aplicadas*, 17, 11-27.
- Casas Gómez, M. y Marín Honor, C. (2021). INTERLEX: Interfax gráfica terminológica. Patente registrada con número 202199904708585, expediente CA-00110-2021 y número de asiento registral 04/2021/4185.
- Cerdà Massó, R. (1986). *Diccionario de lingüística*. Anaya.
- Coseriu, E. (1955-1956). Determinación y entorno. Dos problemas de la lingüística del hablar. *Romanistisches Jahrbuch* 7, 29-54.
- Crespo Miguel, M. (2020). Lingüística digital: revisión de su estado actual y retos en el Instituto Universitario de Investigación en Lingüística Aplicada de la Universidad de Cádiz. *Pragmalingüística*, 28, 148-165.
<https://doi.org/10.25267/Pragmalinguistica.2020.i28.08>

- Díaz Negrillo, A. y Fernández Domínguez, J. (2010). Corpus terminológicos. En Roldán Vendrell, M. (Ed.), Bases para la terminología multilingüe del aceite de olive, olive oil (pp. 37-50). Comares.
- Fernández Smith, G. (2003). Fundamentos teóricos, desarrollo y proyecciones actuales de la lingüística del texto [Tesis doctoral]. Universidad de Cádiz.
- Fernández Smith, G. (2007). Modelos teóricos de la lingüística del texto. Universidad de Cádiz.
- Galina Russell, I. (2011). ¿Qué son las humanidades digitales? *Revista Digital Universitaria*, 12(7), 3-10.
- Gómez Díaz, S. (2022). Lingüística y humanidades digitales: hacia la creación de un repositorio electrónico de documentación lingüística. En R. C. López González (Coord.), *Universalidad y multiversalidad en literatura, lengua y traducción* (pp. 407-413). Comares.
- Gómez Díaz, S. (2024a). Reflexiones lexicográficas en torno al término lingüística del texto. *Pragmalingüística*, 32, 239-262.
- Gómez Díaz, S. (2024b). Delimitación terminológico-conceptual de la lingüística del texto y el análisis del discurso: análisis terminográfico de los usos técnicos y redes conceptuales de estas disciplinas [Tesis doctoral]. Universidad de Cádiz.
- Guerrero Ramos, G. y Pérez Lagos (Coords.) (2021). Terminología, neología y traducción. Comares.
- Harris, Z. S. (1952). Discourse Analysis. *Language*, 28(1), 1-30.
- L'Homme, M. C. (2020). Lexical semantics for terminology: an introduction. John Benjamins.
- Lewandowski, T. (1982). Diccionario de lingüística. Cátedra.
- López Alonso, C. (2014). Análisis del discurso. Síntesis.
- López Yepes, J. A. (2007). ¿Qué es bibliotecología y documentación? *Crónica de un conflicto terminológico y conceptual*. *Bibliotecas*, 25(1), 43-68.
- Macías Couso, A. (2025, en prensa). Tecnología y documentación: aplicaciones al estudio de la historiografía lingüística. En L. Gallardo Calvo y A. C. Romero Vallejo (Eds.), *III Congreso Internacional Tránsitos: Identidades y Culturas en Movimiento*. Trea.

Palek, B. (1968). Cross-Reference. A study from Hyper-Syntax. Charles Universtiy Press.

Pareja, C., Andeyro, Á. y Ojeda, M. (1994). Introducción a la informática. I: Aspectos generales. Complutense.

Ramírez Salado, M. y Álvarez Torres, V. (2018). Métodos y técnicas de detección de unidades terminológicas. En M. Díaz, G. Vaamonde, A. Varela, M. Cabeza Pereiro y J. García-Miguel Gallego (Eds.), Actas do XIII Congreso Internacional de Lingüística Xeral: Vigo, 13-15 de xuño de 2018 (pp. 744-749). Universidad de Vigo.

Rivero Franyutti, A. (2013). ¿Qué son hoy las humanidades y cuál ha sido su valor en la universidad? Revista de la Educación Superior, 42(167), 81-100.

Rodríguez Díez, B. (1981). Las lenguas especiales. El léxico del ciclismo. Colegio Universitario de León.

TeresIA. (2024-). *TeresIA: Portal de acceso a terminologías en España y servicios de inteligencia artificial*. Plataforma Temática Interdisciplinar ES CIENCIA. proyectoteresia.org

TermCat. (1985-). TermCat. *Centre de Terminologia*. termcat.cat

Anexos

Anexo 1. Fichas terminológicas de *jerga*₁ y *lingüística del texto*₇ basadas en el modelo DOCUTERM y realizadas con la plataforma GETERDOC

Término de entrada	JERGA
Subíndice	<i>Jerga</i> ₁
Categoría lingüística del término	Sustantivo
Fuente del término	Abraham, W. (1981). <i>Diccionario de Terminología Lingüística Actual</i> . Madrid: Gredos, 265.
Estatus del término	Documentado en diccionarios y textos, normalizado.
Variación denominativa	-
Área(s) temática(s)	Variación lingüística, sociolingüística, lenguajes especiales, lenguajes marginales (ver notas ₁).

Etimología	Del francés <i>jargon</i> ‘habla incomprensible’.
Mecanismo de formación del término	Préstamo del francés.
Definición	Lenguajes formados en un círculo lingüístico especial.
Fuentes doctrinales de la definición	Abraham, W. (1981). <i>Diccionario de Terminología Lingüística Actual</i> . Madrid: Gredos, 265.
Nivel de metalengua del uso terminológico	General
Contexto(s) representativo(s) de cada uso terminológico	<ol style="list-style-type: none"> 1. No toda variación que dimane de estas u otras causas es, sin embargo, susceptible de ser integrada bajo el concepto «jerga», porque, para dar ese salto cualitativo, se requieren otros condicionamientos: situarse al margen de la variedad estándar, con lo que ello implica de desviación de la norma lingüística, y constituirse en un grupo que diferencie e incluso proteja a sus miembros. La amplitud y ambigüedad de esta caracterización tiene sus ventajas, pero también el inconveniente de dar cobertura al lenguaje del hampa, a los lenguajes especiales de ciertas profesiones y grupos sociales e incluso podría extenderse al habla del pueblo bajo. 2. Estricto sensu, la jerga es una lengua “especial” de un “grupo social determinado” (ej.: lengua de los balleneros Eoki-Foki, lengua de la Matita, de los gitanos; lenguas de ciertas profesiones y de ciertos medios: la tixileira de los “conqueiros”, al suroeste de Asturias; la xiriga de tejeros, canteros y “goxeros” (fabricantes de maconas) de Lianes y Peñamellera; el bron de los caldereros de Miranda (Avilés), el barayete de los afiladores orensanos, etc.)
Fuente de cada contexto	<ol style="list-style-type: none"> 1. Montero Cartelle, E. (2000). El tabú, el eufemismo y las hablas jergales. En Alvar M. (dir.), <i>Introducción a la lingüística española</i>. Barcelona: Ariel, 555 2. Valle, E. R. (1965): ¿Dialecto o germanía?. <i>Español Actual</i>, 5, p. 5
Ejemplos en diferentes lenguas	<ol style="list-style-type: none"> 1. De la jerga deportiva (español): ‘pentatlón’, “conjunto de cinco pruebas atléticas”; ‘swing’, “golpe lateral en el boxeo”. 2. De la jerga de los estudiantes (inglés): ‘beat’, “estar cansado”; ‘wheels’, “automóvil”.
Remisiones a usos sinonímicos o parasinonímicos de carácter terminológico	Relaciones de sinonimia: <i>argot3</i>
Remisiones a usos sinonímicos o parasinonímicos de carácter terminológico	Relaciones de hiperonimia: <i>argot1</i>
Remisiones a usos antonímicos de carácter terminológico	-
Establecimiento de coincidencias terminológicas	-
Equivalencias en otras lenguas	<ol style="list-style-type: none"> 1. Inglés: <i>cant</i> 2. Inglés: <i>slang</i>

	3. Francés: <i>jargon</i>
Fuentes de estas equivalencias	<ol style="list-style-type: none"> 1. VV. AA. (1998). <i>Diccionario terminológico: Lingüística</i>. Barcelona: Larousse, 28. 2. VV. AA. (1998). <i>Diccionario terminológico: Lingüística</i>. Larousse, 28. 3. Marouzeau, J. (1969). <i>Lexique de la terminologie linguistique. Français, Allemand, Anglais-Italien</i>. Paris, Librairie Orientaliste Paul Geuthner, 129. (ver notas₂)
Fuentes documentales del término	<p>Abad Nebot, F. (1986). <i>Diccionario de lingüística de la Escuela Española</i>. Madrid: Gredos, 155.</p> <p>Abraham, W. (1981). <i>Diccionario de Terminología Lingüística Actual</i>. Madrid: Gredos, 265.</p> <p>Alcaraz Varó, E. y Martínez Linares, M. A. (1997). <i>Diccionario de lingüística moderna</i>. Barcelona: Ariel, 312 .</p> <p>Blecua Perdiges, J. M. (1969). <i>Diccionario enciclopédico Salvat Universal. Tomo XIII</i>. Barcelona: Salvat Editores, 182.</p> <p>Cardona, G. R. (1991). <i>Diccionario de lingüística</i>. Barcelona: Ariel, 159.</p> <p>Cerdà Massó, R. (1986). <i>Diccionario de lingüística</i>, Madrid: Anaya, 166.</p> <p>Dubois, J. (1979). <i>Diccionario de lingüística</i>, Madrid: Alianza, 370.</p> <p>Ducrot, O. y Schaeffer, J. M. (1998). <i>Nuevo diccionario enciclopédico de las ciencias del lenguaje</i>. Madrid: Arrecife D. L., 128.</p> <p>Ducrot, O. y Todorov, T. (1983). <i>Diccionario enciclopédico de las ciencias del lenguaje</i>. Madrid: Siglo Veintiuno, 76</p> <p>VV. AA. (1998): <i>Diccionario terminológico: Lingüística</i>. Barcelona: Larousse, 28.</p> <p>Lázaro Carreter, F. (1974). <i>Diccionario de términos filológicos</i>- Madrid: Gredos, 251-252.</p> <p>Luna Traill, E., Viguera Ávila, A. y Báez Pinal, G. E. (2005). <i>Diccionario básico de lingüística</i>, México: Universidad Autónoma de México, 125.</p> <p>Marouzeau, J. (1969). <i>Lexique de la terminologie linguistique. Français, Allemand, Anglais-Italien</i>. Paris: Librairie Orientaliste Paul Geuthner, 28.</p> <p>Martínez de Sousa, J. (1995). <i>Diccionario de Lexicografía Práctica</i>, Barcelona: Vox, 219.</p> <p>Real Academia Española (2014): <i>Diccionario de la lengua española</i> (23ª edición), recuperado de https://dle.rae.es/?id=MQ2t2nA</p> <p>Richards J. C., Platt, J. y Platt, H. (1997). <i>Diccionario de lingüística aplicada y enseñanza de lenguas</i>, Barcelona: Ariel, 229.</p>
Bibliografía recomendada	<p>Casares, J. (1948). <i>Ante el proyecto de un diccionario histórico</i>. Madrid: Imprenta Silverio Aguirre, 32-37.</p> <p>Casares, J. (1969). <i>Introducción a la lexicografía moderna</i>. Madrid: Consejo Superior de Investigaciones Científicas, 278-294.</p> <p>Casas Gómez, M. (1997). <i>Formaciones jergales en el lenguaje de la prostitución. Documentos de trabajo Lynx</i>, 12. Centro de Estudios sobre Comunicación Interlingüística e Intercultural, Valencia: Universidad de Valencia.</p>

	<p>Clavería, C. (1941). Sobre el estudio del argot y del lenguaje popular. <i>Revista Nacional de Educación</i>, 12. Madrid: Centro de Publicaciones del Ministerio de Educación Nacional, 65-80.</p> <p>Clavería, C. (1953). Algunas denominaciones jergales del dinero. <i>Correo Erudito. Gaceta de las Letras y de las Artes</i>, 5(35), 234-235.</p> <p>Lázaro Carreter, F. (1979). Una jerga juvenil: el cheli. <i>Los domingos de ABC (suplemento)</i>. Madrid: ABC, 6.</p> <p>Montero Cartelle, E. (2000). El tabú, el eufemismo y las hablas jergales. En Alvar, M. (dir.), <i>Introducción a la lingüística española</i>. Barcelona: Ariel, 547-563.</p> <p>Suárez Solis, S. (1969). <i>El léxico de Camilo José Cela</i>. Madrid: Alfaguara, pp. 68-73.</p> <p>Torres Martínez, J. C. (1974). "Lógos kryptós" y "akribia" en la función del signo lingüístico. <i>Revista Española de Lingüística</i>, 4(2), 411- 432.</p> <p>Valle, E. R. (1965). ¿Dialecto o germanía? <i>Español Actual</i>, 5, 5-7.</p>
Observaciones y notas	<p>1. El uso de jerga que documentamos aquí es muy abarcador y podemos situarlo dentro de casi cualquier área temática relacionada con la variación lingüística. Teniendo en cuenta que en su definición no se especifica a qué círculo lingüístico especial pertenece, tiene sentido pensar que los factores que intervienen pueden ser múltiples: geográficos, sociales, profesionales, estilísticos, temporales, etc.</p> <p>2. En la definición que Marouzeau proporciona de argot (1968: 28) también aparecen como equivalencias en otros idiomas Sondersprache, Rotwelsch (alemán), slang y cant (inglés) y gergo (italiano). Como remisiones a este último (1968: 265), el autor señala argot y jergon indistintamente, lo que no ocurre con las equivalencias en alemán e inglés, una prueba más de todo lo que abarca este uso de jerga, cuyos límites no están claros. Esto puede ser una de las causas por las que algunos usos de argot y jerga presentan un estado de sinonimia, a pesar de que el primero presenta diferencias con el segundo.</p>
Información en la gestión de la base de datos	<p>Autor(a) de la ficha: Alba Macías Couso. Fecha: mayo de 2025.</p>

Término de entrada	Lingüística del texto
Subíndice	<i>Lingüística del texto</i> ₇
Categoría lingüística del término	Compuesto sintagmático
Fuente del término	Subcorpus sobre lingüística textual y discursiva
Estatus del término	Estandarizado. No documentado lexicográficamente
Variación denominativa	Lingüística textual, LT
Área(s) temática(s)	Lingüística del texto, Análisis del discurso
Etimología	-.

Mecanismo de formación del término	-
Definición	La lingüística del texto es una rama de la Lingüística interesada por la tipología textual y la confección de textos, principalmente los escritos.
Fuentes doctrinales de la definición	Definición propia
Nivel de metalengua del uso terminológico	General
Contexto(s) representativo(s) de cada uso terminológico	<ol style="list-style-type: none"> 1. “En primera instancia, existe un consenso generalizado respecto a que la lingüística textual estudia los textos escritos”. 2. “Se hace especial hincapié en señalar que la disciplina que estudia los discursos orales y escritos es el análisis del discurso, mientras que algunos investigadores prefieren lingüística textual para el discurso escrito”.
Fuente de cada contexto	<ol style="list-style-type: none"> 1. Loureda, Ó. y Schrott, A. (2021): <i>Manual de lingüística del hablar</i>, Berlín: De Gruyter. 2. Fernández Smith, G. (2003): <i>Fundamentos teóricos, desarrollo y proyecciones actuales de la lingüística del texto</i>. Tesis Doctoral, Cádiz: Universidad de Cádiz.
Ejemplos en diferentes lenguas	-
Remisiones a usos sinonímicos o parasinonímicos de carácter terminológico	-
Remisiones a usos hiperónimicos / hiponímicos de carácter terminológico	Hipónimo de <i>Lingüística del texto</i> ³
Remisiones a usos antonímicos de carácter terminológico	<i>Análisis del discurso</i> ⁷
Establecimiento de coincidencias terminológicas	-
Equivalencias en otras lenguas	<ol style="list-style-type: none"> 1. Inglés: <i>textual linguistics</i> 2. Francés: <i>Linguistique textuelle</i>
Fuentes de estas equivalencias	<ol style="list-style-type: none"> 1. House, J. (2012): “Text linguistics and translation”, <i>Handbook of translation studies</i>, 3, pp. 178-183. 2. Adam, J.-M. (1990): <i>Eléments de linguistique textuelle: théorie et pratique de l'analyse textuelle</i>. Liège: Editions Mardaga.
Fuentes documentales del término	No consignado en ninguno de los diccionarios de Lingüística consultados.
Bibliografía recomendada	Koch, P. y Oesterreicher, W. (2007): <i>Lengua hablada en la Rumania: español, francés, italiano</i> . Madrid: Gredos.
Observaciones y notas	Establecemos en este uso una relación de antonimia con respecto a <i>Análisis del discurso</i> ⁷ por oponerse en cuanto al medio de producción de los textos/discursos orales o escritos.

Información en la gestión de la base de datos

Autor(a) de la ficha: Sara Gómez Díaz.
Fecha: septiembre de 2024.

Notas

[1] <https://www.termcat.cat/ca/cercaterm>

[2] <https://proyectoteresia.org/>

[3] Es importante reseñar que este repositorio lingüístico cuenta con antecedentes que se remontan a la recopilación de material bibliográfico iniciado en 1964 con los fondos del Instituto de Miguel de Cervantes del CSIC y, a partir de 1976, con los fondos de Lingüística General y Lengua Española de la Biblioteca Nacional de Madrid por parte del profesor Casas Gómez (Gómez Díaz, 2022).

[4] Una descripción detallada del proceso de digitalización del corpus puede consultarse en Gómez Díaz (2022) y Álvarez Torres y Gómez Díaz (2023). En estas dos publicaciones, las autoras describen la metodología seguida para la confección de este corpus a través de las siguientes fases: 1) digitalización del material lingüístico, 2) tratamiento, conservación y mejora de la calidad de los documentos, 3) almacenamiento y registro de documentos en el repositorio y 4) reconocimiento óptico de caracteres a cada documento para su posterior conversión a formato de texto plano.

[5] A modo de ejemplo, en el apartado «anexos» se mostrará una ficha terminológica del concepto *lingüística del texto* y otra de *jerga*.

[6] Incluimos el símbolo \simeq para identificar la sinonimia y $=$ para las coincidencias terminológicas.

[7] Alemán, árabe, asturiano, catalán, checo, esloveno, español, euskera, finlandés, francés, gallego, griego, inglés, italiano, japonés, ladino, latín, portugués, romaní, rumano, ruso y valenciano.

Nota de los autores

Esta publicación es parte del proyecto de I+D+i ayuda PID2022-139201OB-I00, financiado/a por MCIN/AEI/10.13039/501100011033.

Nota del editor

El editor responsable por la publicación de este trabajo es Isabel Santamaría.

Nota de contribución autoral

El trabajo ha sido realizado de forma equitativa por Sara Gómez Díaz y Alba Macías Couso. Cada investigadora ha contribuido a la configuración y redacción de todos los apartados, si bien el apartado específico sobre las relaciones conceptuales entre *jerga* y *argot* estuvo a cargo de Alba Macías Couso y el apartado sobre los usos terminológicos de *lingüística del texto* estuvo a cargo de Sara Gómez Díaz.

Nota de disponibilidad de datos

El conjunto de datos que apoya los resultados de este estudio no se encuentran disponibles.