

# PROCESAMIENTO DEL LENGUAJE NATURAL

Curso 2017/2018

(Código: 24409117)

## 1. PRESENTACIÓN

La convergencia entre la lingüística y la informática ha resultado en un nuevo perfil profesional e investigador que requiere una formación específica. En sentido estricto, cualquier actividad que implique un análisis o generación de la lengua utilizando el ordenador puede considerarse como procesamiento del lenguaje natural (PLN). No obstante, mientras los lingüistas prefieren hablar de la lingüística computacional como un área de conocimiento dentro de la lingüística aplicada, los informáticos consideran el procesamiento del lenguaje natural como una rama de la inteligencia artificial debido a la posibilidad de desarrollar sistemas de computación que simulen algún aspecto de la capacidad lingüística del ser humano. En este contexto, Procesamiento del Lenguaje Natural es una asignatura teórico-práctica dirigida principalmente a lingüistas y que se centra en dos grandes áreas temáticas:

1. Los fundamentos teóricos básicos del enfoque simbólico en la modelización del conocimiento lingüístico para la construcción de sistemas computacionales que analicen y/o generen textos en lenguaje natural;
2. Las principales aplicaciones de este campo de investigación que ayuden a derribar las barreras comunicativas en la sociedad multilingüe de la información y el conocimiento.

## 2. CONTEXTUALIZACIÓN

*Procesamiento del Lenguaje Natural* es una asignatura de 5 créditos ECTS que pertenece al módulo de contenidos formativos propios del máster *Las TIC en la enseñanza y el tratamiento de las lenguas*.

Debido a su carácter interdisciplinar, esta asignatura resulta igualmente importante tanto para la especialidad (a) *Las TIC para la Enseñanza de Lenguas* como para la especialidad (b) *Las TIC para el Tratamiento de Lenguas*:

- Con respecto a (a), esta asignatura puede ayudar a aquellos alumnos que deseen participar en proyectos de diseño y desarrollo de sistemas inteligentes para la enseñanza de lenguas asistida por ordenador, donde se aplican técnicas propias del procesamiento del lenguaje natural.

- Con respecto a (b), esta asignatura sirve también para ofrecer un acercamiento desde el enfoque computacional a numerosas asignaturas de este máster. Por ejemplo, el tema 2 "Recursos lingüísticos para el procesamiento del lenguaje natural" se relaciona con *La red semántica*, además de con *El análisis de la lengua mediante corpus: técnicas y recursos* dentro del módulo de contenidos comunes. En cuanto a las aplicaciones del procesamiento del lenguaje natural, el tema 4.1 "Traducción automática" se relaciona con *Traducción automática y traducción asistida por ordenador* y el tema 4.3 "Sistemas de diálogo" con *Fonética y tecnologías del habla*.

Procesamiento del Lenguaje Natural no pretende que los alumnos desarrollen aplicaciones informáticas, sino que trata de dotar a los lingüistas del conocimiento necesario para empezar a trabajar conjuntamente con informáticos en proyectos de ingeniería lingüística. En general, esta asignatura posibilita la formación tanto de profesionales que deseen

organizar y coordinar proyectos lingüísticos dentro del marco de las tecnologías del lenguaje, como de investigadores que deseen realizar avances científicos en alguno de los aspectos teóricos del procesamiento del lenguaje natural. Además, permite a los estudiantes conocer a fondo FunGramKB ([www.fungramkb.com](http://www.fungramkb.com)), una base de conocimiento léxico-conceptual diseñada para tareas de PLN, en especial aquellas relacionadas con el procesamiento de texto, como la traducción automática, la extracción y recuperación de información, etc.

### 3. REQUISITOS PREVIOS RECOMENDABLES

A pesar de que *Procesamiento del Lenguaje Natural* no está orientada hacia el tratamiento informatizado de una lengua concreta, el alumno deberá ser capaz de leer y comprender con relativa facilidad textos en inglés, ya que muchas de las lecturas obligatorias propuestas en esta asignatura están escritas en esta lengua.

Igualmente, es muy recomendable que el alumno posea unos sólidos conocimientos sobre aspectos básicos de lingüística teórica, en concreto sobre semántica y sintaxis.

Por otra parte, esta asignatura no exige al alumno conocimiento alguno en programación informática, aunque se le iniciará en la representación formal del conocimiento.

### 4. RESULTADOS DE APRENDIZAJE

Al finalizar el curso, el alumno será capaz de:

1. Analizar los principales logros del procesamiento del lenguaje natural.
2. Identificar y exponer los principales desafíos a los que se enfrenta actualmente el investigador en el procesamiento del lenguaje natural.
3. Evaluar el trabajo cooperativo entre lingüistas e informáticos en el campo del procesamiento del lenguaje natural.
4. Justificar el papel del lingüista en un proyecto de ingeniería lingüística.
5. Describir la formación teórica que precisa el investigador en procesamiento del lenguaje natural.
6. Reflexionar de forma crítica sobre el contenido de artículos científicos relacionados con el procesamiento del lenguaje natural.
7. Redactar un proyecto de investigación en el campo de la ingeniería lingüística

### 5. CONTENIDOS DE LA ASIGNATURA

1. Introducción al procesamiento del lenguaje natural
2. Recursos lingüísticos para el procesamiento del lenguaje natural
3. Tratamiento computacional del contenido lingüístico
  - 3.1 Sintaxis
  - 3.2 Semántica
4. Aplicaciones del procesamiento del lenguaje natural
  - 4.1 Traducción automática
  - 4.2 Recuperación y extracción de información
  - 4.3 Sistemas de diálogo

### 6. EQUIPO DOCENTE

- [ANSELMO PEÑAS PADILLA](#)
- [ALBERTO PEREZ GARCIA-PLAZA](#)

## 7.METODOLOGÍA

La asignatura adoptará una metodología de enseñanza a distancia que garantice la interacción entre el profesor y los alumnos, y favorezca el proceso de aprendizaje autónomo. Con el fin de consolidar el proceso de aprendizaje del alumno, esta metodología se centrará tanto en el trabajo sobre los contenidos teóricos obtenidos a través de la lectura de artículos científicos, como en el desarrollo de actividades prácticas basadas en dichas lecturas. Más concretamente, hemos asignado a cada tema de la asignatura una serie de lecturas obligatorias que presentan con claridad los contenidos teórico-prácticos sobre el procesamiento lingüístico y sus aplicaciones informáticas. A partir de estas lecturas, el alumno demostrará la asimilación de los contenidos mediante la elaboración de tres prácticas y la redacción de un proyecto personal. A continuación, describimos brevemente estos tipos de prueba:

- Las prácticas ofrecerán al alumno la oportunidad de participar en tareas específicas diseñadas a partir de un escenario real de ingeniería del conocimiento. Éstas permitirán a los alumnos demostrar sus destrezas en el campo del tratamiento computacional de la sintaxis y la semántica.

- El proyecto personal consistirá en la redacción de un trabajo que no sólo estimule la reflexión crítica sobre el material estudiado, sino que además exija que el alumno tome decisiones motivadas, anticipe problemas potenciales y proponga posibles soluciones desde la perspectiva del investigador en procesamiento del lenguaje natural.

Los detalles de estas tareas se presentarán en la Guía de Estudio II de la asignatura, que estará disponible al comienzo del curso en la plataforma virtual ALF.

## 8.BIBLIOGRAFÍA BÁSICA

Comentarios y anexos:

En la plataforma virtual, donde el alumno tendrá acceso a los siguientes materiales básicos, se indicará el orden en que deben realizarse las lecturas con el propósito de facilitar la adquisición gradual de los conocimientos requeridos:

Cole, R.A., J. Mariani, H. Uszkoreit, A. Zaenen y V. Zue, eds. (1997) *Survey of the State of the Art in Human Language Technology*. Cambridge: Cambridge University Press.

Gómez Guinovart, X. (2000) "Perspectivas de la lingüística computacional". *Novática: Revista de la Asociación de Técnicos de Informática*, número especial del 25 aniversario (Horizonte 2025), 145, 85-87.

Liddy, E. D. (2001) "Natural language processing". *Encyclopedia of Library and Information Science*. Nueva York: Marcel Decker.

Llisterri, J. (2003) "Lingüística y tecnologías del lenguaje". *Lynx. Panorámica de Estudios Lingüísticos* 2, 9-71.

Mairal Usón, R. y F. Cortés Rodríguez (2006) "An overview of Role and Reference Grammar". En R. Mairal, M. A. Escobar Álvarez, M. S. Peña Cervel, E. Samaniego Fernández y F. Cortés Rodríguez (eds.) *Current Trends in Linguistic Theory*. Madrid: Universidad Nacional de Educación a Distancia, 97-176.

Moreno, L. y A. Molina. (1999) "Preliminares y tendencias en el Procesamiento del

Lenguaje Natural". *Revista de Inteligencia Artificial* 99 (7).

Moure, T. y J. Llisterri (1996) "Lenguaje y nuevas tecnologías. El campo de la lingüística computacional". En M. Fernández Pérez (ed.) *Avances en lingüística aplicada*. Santiago de Compostela: Universidade de Santiago de Compostela, 147-228.

Nolan, B. and Y. Salem (2011) "UniArab: RRG Arabic-to-English Machine Translation." En Watara Nakamura (ed.) *New Perspectives in Role and Reference Grammar*. London: Cambridge Scholars Publishing. 312-344.

Periñán Pascual, C. (2005) "Procesamiento del lenguaje natural: de lingüista a ingeniero del conocimiento". En I. K. Brady, M. Navarro Coy y C. Periñán Pascual (eds.) *Nuevas Tendencias en Lingüística Aplicada*. Murcia: Quaderna, 293-317.

Periñán Pascual, C. y F. Arcas Túnez (2010) "The architecture of FunGramKB". *Proceedings of the 7th International Conference on Language Resources and Evaluation*. European Language Resources Association (ELRA). Valletta (Malta), 2667-2674.

Periñán Pascual, C. y R. Mairal Usón (2009) "Bringing Role and Reference Grammar to natural language understanding". *Procesamiento del Lenguaje Natural* 43, 265-273.

Periñán Pascual, C. y R. Mairal Usón (2010) "Enhancing UniArab with FunGramKB". *Procesamiento del Lenguaje Natural* 44, 19-26.

Periñán Pascual, C. y R. Mairal Usón (2010) "La gramática de COREL: un lenguaje de representación conceptual". *Onomázein* 21, 11-45.

Rodríguez, H. (2000) "Técnicas básicas en el tratamiento informático de la lengua". *Quark. Ciencia, Medicina, Comunicación y Cultura* 19: 26-34.

## 9. BIBLIOGRAFÍA COMPLEMENTARIA

Comentarios y anexos:

Dale, R., H. Moisl y H. Somers (eds.) (2000): *Handbook of Natural Language Processing*, Nueva York, Marcel Dekker.

Mitkov, R. (ed.) (2003): *The Oxford Handbook of Computational Linguistics*, Oxford, Oxford University Press.

## 10. RECURSOS DE APOYO AL ESTUDIO

La plataforma de aprendizaje en Internet permitirá no solo la interacción entre los alumnos y el profesor, sino también el acceso a las lecturas obligatorias y a las instrucciones para la realización de los trabajos propuestos. Igualmente, el foro de debate y el correo electrónico se convierten en instrumentos imprescindibles para la consulta y la resolución de dudas relacionadas con el trabajo del alumno.

## 11. TUTORIZACIÓN Y SEGUIMIENTO

El alumno podrá contactar con el profesor responsable de la asignatura a través de los siguientes medios:

- Correo electrónico: [pedrou@invi.uned.es](mailto:pedrou@invi.uned.es)

El horario de atención telefónica al alumno se indicará en la plataforma virtual a principio del curso.

## 12.EVALUACIÓN DE LOS APRENDIZAJES

La evaluación será continua. La calificación final se obtendrá de la suma de las notas obtenidas en la primera práctica (20%), la segunda práctica (20%), la tercera práctica (20%) y el proyecto personal (40%). Será necesario entregar cada prueba en el plazo que se solicite. La descripción detallada de las pruebas, los plazos de entrega y los criterios de evaluación se presentarán en la plataforma virtual a principio del curso.

## 13.COLABORADORES DOCENTES

Véase equipo docente.

## 14.Opiniones de antiguos alumnos

“Es de agradecer la clarísima organización de la guía de estudio...¡Es imposible perderse! Se ha abierto un campo hasta ahora absolutamente desconocido para mí...y realmente se ha producido el acercamiento a FunGramKB.” (Teresa Llopis Guixot, curso 2012-2013)

“¡Esta asignatura es un reto muy interesante!” (Alicia Barra Salinas, curso 2012-2013)

“El nivel y la generosidad en las respuestas que hemos tenido en este foro supera con mucho los de otras materias. ¡Hacía tiempo que no lo pasaba tan bien estudiando algo!” (Jordi Viña Domingo, curso 2012-2013)

“Realmente ha sido una de las asignaturas que más he disfrutado en lo que va del Máster. Muy bien planificada y con actividades pertinentes y significativas” (Soledad Álvarez, curso 2013-2014)

“Personalmente el proyecto final ha sido lo mejor para mí, ya que es cuando he entendido todos los conceptos abstractos que hemos trabajado y dónde encajan en un proceso de PLN” (Gorka González González, curso 2013-2014)

“La verdad es que PLN es una asignatura compleja a la que he tenido que dedicar mucho tiempo. De hecho, la cogí con muchas ganas precisamente porque me suponía un reto, un reto que al final he superado, lo que me hace sentir muy feliz. Ahora bien, nada de esto habría sido posible sin los constantes mensajes de ánimo de la profesora, sin sus explicaciones, sin su paciencia...” (María García Pena, curso 2013-2014)