23-24

MÁSTER UNIVERSITARIO EN CIENCIA DEL LENGUAJE Y LINGÜÍSTICA HISPÁNICA

GUÍA DE ESTUDIO PÚBLICA

LINGÜÍSTICA COMPUTACIONAL

CÓDIGO 24402743

UNED

23-24

LINGÜÍSTICA COMPUTACIONAL

CÓDIGO 24402743

ÍNDICE

PRESENTACIÓN Y CONTEXTUALIZACIÓN
REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA ASIGNATURA
EQUIPO DOCENTE
HORARIO DE ATENCIÓN AL ESTUDIANTE
COMPETENCIAS QUE ADQUIERE EL ESTUDIANTE
RESULTADOS DE APRENDIZAJE
CONTENIDOS
METODOLOGÍA
SISTEMA DE EVALUACIÓN
BIBLIOGRAFÍA BÁSICA
BIBLIOGRAFÍA COMPLEMENTARIA
RECURSOS DE APOYO Y WEBGRAFÍA
IGUALDAD DE GÉNERO

Nombre de la asignatura	LINGÜÍSTICA COMPUTACIONAL
Código	24402743
Curso académico	2023/2024
Título en que se imparte	MÁSTER UNIVERSITARIO EN CIENCIA DEL LENGUAJE Y LINGÜÍSTICA HISPÁNICA
Tipo	CONTENIDOS
Nº ETCS	5
Horas	125.0
Periodo	SEMESTRE 2
Idiomas en que se imparte	CASTELLANO

PRESENTACIÓN Y CONTEXTUALIZACIÓN

El objetivo principal de esta asignatura es abordar los problemas computacionales del procesamiento lingüístico, fundamentalmente desde un punto de vista teórico, aunque sin renunciar a que el alumno tenga una experiencia práctica en el uso de herramientas y recursos habituales en la lingüística computacional.

La lingüística computacional aporta a los estudios lingüísticos el grado de formalización necesario para poder considerarlos científicos. En efecto, buena parte de los trabajos en lingüística teórica, incluso realizados en el marco de escuelas con un fuerte aparato teórico, se basan más en la argumentación que en la formalización y no resisten una implementación del modelo propuesto. La lingüística computacional se apoya en tales teorías lingüísticas para, con ayuda de los formalismos gramaticales adecuados, representar el conocimiento lingüístico como tal y dar solución a problemas concretos del procesamiento lingüístico. Estas soluciones, estén cognitivamente motivadas o sean simplemente soluciones de ingeniería, permiten a los humanos resolver problemas cotidianos relacionados con el procesamiento lingüístico, y contribuyen a que el alumno descubra un ámbito bien definido de aplicación de sus conocimientos.

La lingüística computacional es una disciplina reciente que se ha incorporado a los planes de estudio de las licenciaturas y grados de formación lingüística. Sin embargo, como otros aspectos de la formalización lingüística, ha estado ausente de los planes de estudio de Filología y Humanidades. Este hecho ha favorecido la incorporación de Ingenieros y licenciados en ramas técnicas con conocimientos de programación al mercado de trabajo de la ingeniería lingüística en detrimento de quienes tienen la lengua como objeto de estudio. Este ámbito de la ingeniería lingüística es, sin duda, uno de los que mayor capacidad de absorción de nuevos licenciados tiene. Baste, para ello, con echar un vistazo a las empresas que desarrollan tecnologías lingüísticas asociadas a la Red.

Sin embargo, sería descabellado pensar que un solo curso de estas características puede formar al alumno en los conocimientos necesarios para su incorporación al mercado de trabajo o al campo de la investigación en el ámbito de las tecnologías lingüísticas. En todo caso, los conocimientos básicos adquiridos y las destrezas que el alumno pueda desarrollar en su trabajo de implementación le servirán de ayuda para valorar el mejor itinerario de (auto)formación, si su deseo es explorar esta línea profesional.

La asignatura se imparte en el segundo semestre del curso. Pertenece al Módulo de Especialidad, dentro de la opción A (*Lenguaje y Comunicación*), al igual que *Lingüística,*

adquisición de la lengua y patologías del lenguaje, Fundamentos de lenguaje visual, *Fundamentos fonológicos e icónicos de la comunicación, Lenguaje y comunicación en los medios* y *Componentes lingüísticos de la traducción.* Complementa y desarrolla los objetivos de formación en competencias, destrezas y habilidades teóricas y prácticas correspondientes a los Módulos de Contenidos Comunes y Contenidos Formativos Propios del Máster en Ciencia del Lenguaje y Lingüística Hispánica.

Presentación del equipo docente

Fernando Sánchez León fue director del Departamento de Tecnología de la Real Academia Española en las áreas de Lingüística Computacional e Ingeniería de Software. Es doctor en Filosofía y Letras por la Universidad Autónoma de Madrid (1997), universidad en la que ha sido profesor de Lingüística General y de Lingüística Computacional. Trabaja desde hace 30 años en lingüística computacional, disciplina en la que se ha formado gracias a su participación en proyectos europeos dedicados a la traducción automática (EUROTRA, 1987-1992), corrección gramatical (GramCheck, 1994-1995) y lingüística de corpus (CRATER, 1994-2005). En los últimos años, ha centrado su actividad en el desarrollo de herramientas y recursos de calidad para el procesamiento lingüístico automático en español, en sus dimensiones sincrónica y diacrónica (CREA y CORDE, 1995-2000), así como en el desarrollo de técnicas de extracción de términos a partir de corpus especializados (CCT, 2003-2006).

REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA ASIGNATURA

- Habida cuenta de la escasa formación que tradicionalmente muestra el alumnado de titulaciones filológicas en tecnología, no se impone ningún requisito especial para cursar esta asignatura. Para la realización de buena parte de los ejercicios y del trabajo basta con que el alumno tenga una actitud crítica y reflexiva sobre los problemas y métodos en lingüística teórica, sin importar demasiado la escuela gramatical. Con todo, buena parte del material supone un cierto conocimiento del generativismo, como modelo de formalización gramatical, por lo que los alumnos que conozcan esta escuela podrán leer entre líneas y sacar mejor provecho a algunas lecturas complementarias.
- El alumno debe tener competencia lectora en inglés, pues una parte importante de la bibliografía obligatoria está redactada en esta lengua.
- Debe tener soltura en el manejo de los ordenadores, a nivel de usuario, con el fin de poder instalar y ejecutar los programas que se incluyen con el material didáctico. En el caso de que surjan problemas o dudas en este particular, el profesor estará a disposición del alumno en el horario de tutorías para guiarlo hasta que pueda moverse con soltura en el manejo de estos programas. Si los problemas afectan al manejo de la plataforma informática de la UNED, hay un teléfono (34 913 988 801) y un servicio de gestión de incidencias (acceso)

para solucionarlos.

•El material del curso da por supuesto un conocimiento básico de los tradicionales niveles de descripción de la gramática y de los problemas que se tratan en cada uno de ellos. Con todo, suponemos que otros cursos de este título de máster contribuirán a reducir las diferencias de saberes de estos titulados y los filólogos.

EQUIPO DOCENTE

Nombre y Apellidos	RICARDO MAIRAL USON
Correo Electrónico	rmairal@flog.uned.es
Teléfono	8645/6819
Facultad	FACULTAD DE FILOLOGÍA
Departamento	FILOLOGÍAS EXTRANJERAS Y SUS LINGÜÍSTICAS

COLABORADORES DOCENTES EXTERNOS

Nombre y Apellidos	FERNANDO SÁNCHEZ LEÓN
Correo Electrónico	fsanchez@invi.uned.es

HORARIO DE ATENCIÓN AL ESTUDIANTE

Los alumnos pueden plantear sus dudas y realizar las consultas que consideren convenientes al profesor de la asignatura, a través preferentemente del correo electrónico: fernando.sanchez.leon@gmail.com.

En última instancia, podrá recurrirse al teléfono (34 915 061 666) o al mensaje dirigido a la dirección postal:

Fernando Sánchez León
c/ Los Nogales, 1, esc. 2, 5.º A
28005 Madrid

El **horario de atención** es:

Martes y jueves, de 16:30 a 18:30 h.

Se entiende que quedan exceptuados los períodos de vacaciones, las semanas de examen, los sábados y festivos -en Madrid- y las fechas (imposibles de prever) en que el profesor asista a reuniones, congresos, etc.

COMPETENCIAS QUE ADQUIERE EL ESTUDIANTE

Competencias Básicas:

CB6 - Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación

CB7 - Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio

CB8 - Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

CB9 - Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades

CB10 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

Competencias Generales:

2.1.1. - Comunicación y expresión escrita

2.1.2. - Comunicación y expresión oral

2.1.3. - Comunicación y expresión en otras lenguas

3.2. - Habilidad para negociar de forma eficaz

3.3. - Habilidad para la mediación y resolución de conflictos

3.4. - Habilidad para coordinar grupos de trabajo

3.5. - Liderazgo

4.1. - Conocimiento y práctica de las reglas del trabajo académico

4.2. - Compromiso ético y ética profesional

4.3. - Conocimiento, respeto y fomento de los valores fundamentales de las sociedades democráticas

2.1.4. - Comunicación y expresión matemática, científica y tecnológica

2.2.1. - Competencia en el uso de las TIC

2.2.2. - Competencia en la búsqueda de la información relevante

2.2.3. - Competencia en la gestión y organización de la información

2.2.4. - Competencia en la recolección de datos, el manejo de bases de datos y su presentación

3.1. - Habilidad para coordinarse con el trabajo de otros

Competencias Específicas:

CE1 - Que el estudiante adquiera una alta capacitación académica en medios relacionados con el lenguaje, la comprensión, expresión, comunicación y ejercicio crítico del pensamiento. Los contenidos y métodos seguidos dotan al estudiante de conocimientos más específicos del lenguaje que los adquiridos en el Grado y lo disponen a la investigación, preparación de

proyectos, comprensión de otras áreas científicas a través del análisis lingüístico, elaboración de Tesis doctorales en posibles campos de preferencia suya como el estudio, por ejemplo, de las relaciones del lenguaje con el genoma, los conocimientos cibernéticos de la comunicación, la investigación etnolingüística, la importancia del discurso expresivo en los medios comunicativos y la proliferación de modelos gramaticales, etc.
CE2 - Adquirir una formación más directa entre las relaciones del lenguaje y la imagen en diferentes niveles del discurso, sus implicaciones mixtas en campos como el de la escritura, funcional y poética, la formación del contenido lingüístico según teorías de modelos actuales de la tecnociencia, o una aún más específica de la lingüística propiamente hispana, que atiende al orden diacrónico, sincrónico, expresivo y de sistematización gramatical propia, en comparación con otros dominios de la ciencia actual.
CE3 - Adquirir una alta capacitación académica para iniciar estudios sólidos o proyectos de investigación concernientes al Lenguaje o relaciones lingüísticas de campos icónicos y conceptuales como los de Antropología, Poética, Medios informativos, Filosofía fenomenológica, Hermenéutica, Biogenética, etc. Asimismo, adquirir la capacitación para el ejercicio no sólo profesional en áreas docentes o técnicas relacionadas con el Lenguaje, sino también con el mundo de la empresa por la capacidad de síntesis y análisis que aporta, así como de “economía” de medios en relación con el balance de “producciones” posibles y “rendimientos” de un programa específico: concepción, organización, planificación, interactividad comunicativa, etc.
CE4 - Adquirir una avanzada capacitación para el análisis rápido de nuevos entornos sociovitales y profesionales, sobre todo en aquellas áreas que requieren una alta capacidad de adaptación inmediata, como las relacionadas con los medios de comunicación, el estudio comprehensivo, la política, publicidad, empresa, etc.
CE5 - Adquirir el grado de autonomía crítica necesario para saber exponer con rigor conceptual y medios críticos razonados cualquier tema y derivaciones suyas conceptuales, como objetivo máximo del estudio humanístico.
CE6 - Adquirir las competencias como experto en Lingüística y que esta figura adquiera crédito social además de garantía investigadora. Así entendemos la etiqueta de “Académico-mixto”.

RESULTADOS DE APRENDIZAJE

El curso pretende que el estudiante conozca los hitos más relevantes de la lingüística computacional, los problemas que esta plantea y las aplicaciones que se desarrollan en esta área de investigación. Pretende, asimismo, que se familiarice con la formalización de problemas lingüísticos y para ello profundice en el estudio (y en la experimentación) de los formalismos de representación del conocimiento lingüístico más utilizados en la actualidad, resolviendo problemas concretos en varios de los niveles de descripción gramaticales.

El programa de la asignatura persigue la consecución de los siguientes resultados de aprendizaje:

•Que el estudiante revise y comprenda los fundamentos, problemas y métodos de los modelos computacionales del lenguaje dominantes en la actualidad: el probabilístico y el

simbólico, con especial hincapié en el segundo de ellos. En el curso no se aborda el paradigma neuronal, dado que no se supone al alumno los conocimientos matemáticos necesarios para comprender las bases teóricas de este paradigma.

- Que conozca someramente la historia de la lingüística computacional, así como las aplicaciones más reseñables del procesamiento del lenguaje natural.
- Que estudie y, por tanto, conozca en profundidad, los formalismos más usuales para la representación del conocimiento lingüístico en los niveles morfológico, sintáctico y semántico.

CONTENIDOS

BLOQUE I. Introducción

BLOQUE II. Problemas del procesamiento y rudimentos de programación lógica

BLOQUE III. Niveles de descripción gramatical

BLOQUE IV. Anotación lingüística de corpus

METODOLOGÍA

El material del curso se distribuye a partes iguales entre materiales preparados expresamente para esta asignatura y artículos (o capítulos de libros) de interés de prestigiosos investigadores que cubren parcelas de estudio teórico. Los primeros contienen un volumen importante de ejercicios, que el alumno deberá realizar, así como una serie de problemas nuevos que se le plantean sobre los ejercicios ya vistos o los conocimientos recién adquiridos. Se pretende con ello que el alumno desarrolle su capacidad de aprendizaje activo con un conjunto de casos que bien pudieran ser reales en un entorno profesional. Por su parte, para que el profesor tenga constancia de la asimilación de las lecturas del segundo tipo, el alumno deberá remitir al profesor un breve documento resumen de las ideas fundamentales de cada una de estas lecturas, así como un comentario crítico de las mismas. Por tanto, el alumno entregará un resumen por cada una de las lecturas que no contienen ejercicios (debido a que su contenido es, eminentemente teórico). Sin embargo, de aquellas que contienen ejercicios (convenientemente destacados en el texto de la lectura), la resolución de estos será suficiente para valorar los avances del alumno, no siendo necesario enviar, además, resumen de lectura. El orden de lectura y estudio de uno y otro tipo de material se entrelaza, de forma que se favorezca el seguimiento del alumno en la

adquisición de los nuevos conocimientos y habilidades. Finalmente, el alumno debe realizar un trabajo práctico de implementación de una gramática con ayuda de un entorno de software para la creación de gramáticas de unificación. Para ello, el alumno deberá haber entendido los conceptos que subyacen al concepto de unificación y a las jerarquías de rasgos, y haber asimilado los conceptos sobre análisis sintáctico automático que se le presentan en uno de los capítulos del curso.

El alumno debe enfocar el estudio de la asignatura como un proceso continuo y constante, al que debe dedicar, si es posible todos los días, un tiempo razonable. Este enfoque favorecerá la comprensión y la capacidad de reflexión sobre el material que se le entrega. De hecho, algunos materiales son relativamente densos, tanto en la exposición, como por el tema mismo que abordan, lo que requiere una lectura pausada de los mismos y no un intento de asimilación veloz y voraz de los contenidos.

En general, el material no se ha desarrollado para el estudio memorístico sino más bien para la reflexión y el desarrollo de la capacidad de análisis y el razonamiento crítico. A estos objetivos contribuirá un tratamiento pausado y constante del material.

Se recomienda vivamente la lectura del material en el orden que se sugiere, pues la presentación de conceptos y problemas es incremental, por lo que otros itinerarios pueden llevar a la incomprensión y, por ende, al desaliento del alumno.

En la lectura/estudio del material, se recomienda que se sigan las pautas habituales de estudio, de forma que el trabajo con los textos se reduzca a las tareas de:

1. identificación de la estructura lógica del texto, con objeto de ir anticipando las cuestiones que el capítulo plantea;
2. lectura(s) del texto, buscando las ideas principales, los conceptos nuevos, su definición y su relación con los conceptos anteriores, del texto o de toda la asignatura;
3. resumen de las ideas principales del texto, realizando una reflexión, en los términos propios del alumno, de las aportaciones del mismo;
4. asimilación de estas ideas, intentando fijar los conceptos fundamentales, relacionándolos con los conceptos aprendidos anteriormente y reproduciéndolos mentalmente;
5. realización de los ejercicios y resolución de los problemas planteados en todos aquellos capítulos que los incluyen. Unos y otros deberán realizarse después de haber leído y asimilado todos los conceptos que plantea el capítulo correspondiente;
6. relectura del texto: esta segunda lectura permitirá al alumno asimilar la terminología del dominio. Efectivamente, solo tras haber comprendido y asimilado los conceptos nuevos, se está en condiciones de asegurar la adecuada asimilación terminológica, que es, por otra parte, una de las características del discurso vinculado al conocimiento de una materia.

Una vez terminado el estudio de un tema y realizados los ejercicios y resueltos los problemas (si procede), el alumno estará en condiciones de plantear sus dudas al profesor de la asignatura y de hacerle llegar su solución a los problemas o, en su caso, el resumen de ideas fundamentales del tema. El profesor atenderá las dudas del alumno y comentará su solución a los problemas. Se creará una lista de distribución que permita enviar a todos los

alumnos los comentarios a aquellas dudas que puedan resultar de interés general.

Cuando el alumno inicie el estudio del material didáctico correspondiente al capítulo sobre sintaxis, decidirá, junto con el profesor, el tema de trabajo práctico de implementación que deberá realizar como parte de los requisitos de evaluación del curso. En todo momento, por medio de las tutorías, recibirá consejos y comentarios sobre su trabajo hasta la entrega del mismo al final del curso.

SISTEMA DE EVALUACIÓN

TIPO DE PRUEBA PRESENCIAL

Tipo de examen	No hay prueba presencial

CARACTERÍSTICAS DE LA PRUEBA PRESENCIAL Y/O LOS TRABAJOS

Requiere Presencialidad	No

Descripción

No hay prueba presencial. El sistema de evaluación de la asignatura está basado en la presentación de una **monografía final** y **dos pruebas de evaluación continua** (PEC).

MONOGRAFÍA FINAL

La calificación final, además de tener en cuenta el trabajo de los alumnos durante el curso, del que el profesor tendrá información puntual por medio de los ejercicios y los resúmenes de lectura, se fundamentará en el trabajo de diseño, implementación y documentación de una gramática sintáctica de la lengua de elección del alumno; en este trabajo se desarrollará una solución de ingeniería a un problema concreto del nivel sintáctico. Por su parte, la documentación de la gramática deberá tratar tanto los aspectos lingüísticos como computacional, y deberá acompañarse de comentarios sobre cobertura y sobregeneración. Las directrices para la realización del trabajo se encuentran en el material didáctico que se proporciona al alumno.

Criterios de evaluación

El trabajo se califica de 0 a 6 puntos.

Se enviará a la dirección electrónica del profesor de la asignatura.

Ponderación de la prueba presencial y/o los trabajos en la nota final	La monografía final obligatoria supone, al menos, el 60% de la calificación final.
Fecha aproximada de entrega	Coincidiendo con el período de exámenes de la convocatoria de junio. En todo caso, antes de las 23:55 h del viernes de la segunda semana de exámenes (v. el calendario general de la UNED).

Comentarios y observaciones

Es muy conveniente -en previsión de fallos informáticos, incluida la sobrecarga en las líneas o en el servidor de la UNED- no esperar al último momento para cargar el ejercicio.

La calificación de la monografía final de la convocatoria de junio, si no llega al aprobado, no es válida para la convocatoria extraordinaria de septiembre. En esta convocatoria puede volver a presentarse antes de las 23:55 h del viernes inmediatamente posterior al primer lunes del mes.

PRUEBAS DE EVALUACIÓN CONTINUA (PEC)

¿Hay PEC? Si,PEC no presencial

Descripción

La evaluación continua de esta asignatura se basará en la realización de dos pruebas.

Para resolver cualquier duda relativa a la materia impartida, el profesor asesorará a los estudiantes, por correo electrónico.

En el horario que se indica en el apartado de Equipo docente, los alumnos podrán consultar telefónicamente al profesor cualquier asunto que sea de su interés.

PEC 1

Ejercicios propuestos en el material didáctico.

PEC 2

Resumen de lecturas.

Criterios de evaluación

El alumno deberá enviar a la dirección electrónica del profesor todos los ejercicios que se proponen en el material didáctico. Se espera del alumno una solución razonada (dado que todos los ejercicios obligan a realizar un esfuerzo de reflexión y adaptación de ejemplos presentados en el material didáctico) en la que se documenten las decisiones adoptadas. Asimismo, el alumno, tras el estudio de todos los temas que no contengan ejercicios (quizá complementado con la lectura del material complementario), redactará un resumen crítico de las ideas fundamentales de cada artículo que deberá enviar al profesor. Este resumen puede realizarse a modo de presentación en aplicaciones como PowerPoint y, en cualquier caso, deberá enviarse al profesor, como la solución a los ejercicios, por correo electrónico.

Cada una de las PEC tendrá una calificación máxima de 2 puntos.

Ponderación de la PEC en la nota final: Las pruebas de evaluación continua pueden suponer hasta un 40% de la calificación final.

Fecha aproximada de entrega: Durante las semanas 7.ª y 12.ª del curso, aproximadamente. V. el calendario fijado en la página de la asignatura.

Comentarios y observaciones

Estas pruebas solo pueden entregarse en el plazo fijado para ello durante el segundo cuatrimestre mediante envío al correo electrónico del profesor de la asignatura. Es muy conveniente -en previsión de fallos informáticos, incluida la sobrecarga en las líneas o en el servidor de la UNED- no esperar al último momento para cargar el ejercicio.

Las calificaciones de las pruebas de evaluación continua se tendrán en cuenta tanto para la convocatoria ordinaria de junio como para la convocatoria extraordinaria de septiembre. Pero los plazos de presentación solo se abrirán en las fechas arriba indicadas, durante el desarrollo del segundo cuatrimestre (febrero-junio).

OTRAS ACTIVIDADES EVALUABLES

¿Hay otra/s actividad/es evaluable/s? No

Descripción

No hay otras actividades.

Criterios de evaluación

No hay otras actividades.

Ponderación en la nota final No hay otras actividades.

Fecha aproximada de entrega No hay otras actividades.

Comentarios y observaciones

No hay otras actividades.

¿CÓMO SE OBTIENE LA NOTA FINAL?

La calificación final de la asignatura se obtiene ponderando los resultados de la monografía final (hasta un máximo de 6 puntos sobre 10) y los de la evaluación continua desarrollada a lo largo del curso (hasta un máximo de 2 puntos por cada una de las PEC).

BIBLIOGRAFÍA BÁSICA

Lecturas obligatorias

Se indica, con un número entre corchetes, el capítulo del temario al que se refieren. Aquellas referencias que llevan además un asterisco se incluyen en el espacio virtual del curso. La bibliografía, para facilitar la relación con el temario de la asignatura, se indica en el orden de lectura.

•Sánchez León, F. (2008) “Panorama general e histórico, definición(es), aproximación, modelos”. [1, *]

•European Commission (1997?) “Language Engineering: Harnessing the power of language”. [1, *]

•Llisterri, J. (2003) “Las tecnologías del habla: Entre la ingeniería y la lingüística”, en *Actas del Congreso Internacional "La Ciencia ante el Público. Cultura humanística y desarrollo científico y tecnológico"*. Universidad de Salamanca, 28-31 de octubre 2002. Salamanca:

Instituto Universitario de Estudios de la Ciencia y la Tecnología. Edición en CD-ROM. pp. 44-67. [2, *]

•King, M. (1988?) “A General Introduction to Natural Language Processing”. [3, *]

•Sánchez León, F. (2008) “Introducción a Prolog” [4, *].

•Grefenstette, G. and P. Tapanainen (1994) “What is a word, What is a sentence? Problems of tokenization”, in *3rd International Conference on Computational Lexicography*, Budapest, pp. 79-87. [5, *]

•Sánchez León, F. (2008) “Autómatas de estados finitos”. [6, *]

•Sánchez León, F. (2008) “Fonología/Morfología de estados finitos”. [6, *]

•Karttunen, L., J-P. Chanod, G. Grefenstette and A. Schiller (1996) “Regular Expressions for Language Engineering”. *Natural Language Engineering*, 2(4): 305-238. [6, *]

•Sánchez León, F. (2008) “Gramáticas y lenguajes formales”. [7, *]

•Sánchez León, F. (2008) “Introducción a los formalismos gramaticales basados en unificación”. [7, *]

•Sánchez León, F. (2008) “*Parsing*: estrategias”. [7, *]

•McConnel, S. (2002) “PC-PATR Reference Manual: A unification based parser”. [7, *]

•Agirre, E. and Ph. Edmonds (2006) “Introduction”, in *Word Sense Disambiguation: Algorithms and Applications*, Springer. [8, *]

•Sánchez León, F. (2008) “Anotación morfosintáctica de corpus”. [9, *]

Bibliografía recomendada

•Jurafsky D. and J. H. Martin (2000) *Speech and language Processing*. Prentice Hall, New Jersey.

•Allen, J. F. (1994) *Natural Language Understanding*. Addison-Wesley, Reading, MA, 2nd edition.

•Moreno Sandoval, A. (1998) *Lingüística Computacional*. Síntesis, Madrid.

BIBLIOGRAFÍA COMPLEMENTARIA

Bibliografía complementaria comentada

Se presenta únicamente la bibliografía complementaria relativa al bloque 1, pues el resto se ha integrado en los capítulos correspondientes del material didáctico. Como con las lecturas obligatorias, se incluye un asterisco al final de aquellas referencias que se incluyen en el espacio virtual del curso:

•Gómez Guinovart, J. (1998): “Fundamentos de Lingüística Computacional: bases teóricas, líneas de investigación y aplicaciones”. En J. Baró i Queralt y P. Cid Leal (eds.), *Anuari SOCADI de Documentació i Informació*, Barcelona: Societat Catalana de Documentació i Informació, 135-146. Interesante artículo introductorio en el que se presentan los grados de imbricación entre Lingüística e Informática y las distintas líneas de trabajo en función de esa imbricación. Con todo, téngase en cuenta que han transcurrido más de 20 años desde su

publicación y que, por tanto, algunas limitaciones tecnológicas que plantea el artículo han sido superadas. Incluye una sección de fuentes de información relacionadas con esta disciplina. [*]

- Jurafsky, D. y Martin, J. H. (2000): Capítulo 1 “Introduction”. Repaso general a los problemas, modelos y algoritmos para el procesamiento lingüístico, junto con un breve apunte histórico. [*]
- Automatic Language Processing Advisory Committee (1966): *Language and Machines. Computers in Translation and Linguistics*. El alumno interesado en conocer las razones que se adujeron para el descalabro de la Traducción Automática a mediados de los 60, puede encontrar el informe ALPAC en la red http://books.nap.edu/openbook.php?isbn=ARC000005
- Moreno Sandoval, A. (1998): *Lingüística computacional*. Síntesis, Madrid. Se trata de una introducción básica muy asequible; por este motivo se ha incluido en la bibliografía recomendada. No obstante, debe manejarse con cautela, pues, aunque el planteamiento de la obra es adecuado, contiene algunos errores de fondo. El alumno deberá contrastar con el profesor de la asignatura aquellos aspectos contradictorios con las ideas del resto de la documentación del curso.

RECURSOS DE APOYO Y WEBGRAFÍA

El estudiante cuenta con una serie de recursos de apoyo virtual, así como con otros de tipo didáctico que facilitan la preparación de la asignatura conforme a la metodología propia de esta Universidad:

A. Recursos de apoyo en la plataforma de la UNED

La asignatura está virtualizada en la plataforma aLF de la UNED. A través de ella se accede a la mayor parte de los materiales de estudio (en el apartado de *Documentos*).

Para el resto de las actividades debe recurrirse a la dirección electrónica del profesor: fernando.sanchez.leon@gmail.com.

B. Medios didácticos utilizados

La asignatura dispone de:

- Esta **Guía de la asignatura. Para** los alumnos inscritos en la materia, la Guía se complementa, una vez comenzado el curso, con nuevos apartados que ofrecen más información en cuanto a los **contenidos** de la asignatura y al **plan de trabajo** recomendado.
- Una **bibliografía básica** de consulta imprescindible.
- Unas **lecturas complementarias.**
- Una serie de **documentos** que desarrollan parte del programa.

C. Recursos de apoyo

Software para prácticas

Los programas y lenguajes de programación que se utilizarán para ejemplificar aspectos del curso se incluyen también en la carpeta del espacio virtual. Debe hacerse notar que este no es un curso de programación para lingüistas y que, por tanto, será difícil, en la mayoría de los casos, usar algunos de los programas y lenguajes para algo más que probar en casa los ejemplos que se explican en la documentación del curso.

En la elección de programas y lenguajes se ha optado por aquellos que se encuentran en el dominio público (son gratuitos) y que tienen versiones bajo sistema operativo Windows (en cualquiera de sus versiones). Se ha supuesto, por tanto, que el alumno tiene alguna versión de este sistema operativo instalada en su ordenador. Todo el material, sin embargo, se ha desarrollado y probado bajo Linux, sistema operativo para el que, lógicamente, existen versiones de estos mismos programas.

Para instalar los programas incluidos en la carpeta que contiene el material didáctico, el alumno deberá seguir las instrucciones que se incluyen en cada caso. Todos los programas se encuentran en el directorio llamado ‘software’. Todos tienen un desinstalador, por lo que podrán eliminarse del disco duro, si se desea, cuando termine el curso.

A título meramente indicativo, además de un programa de descompresión de archivos, la mencionada carpeta contiene un intérprete de Perl (ActivePerl), un intérprete de Prolog (SWI-Prolog) y el conjunto de programas para el desarrollo de procesadores morfológicos y sintácticos con gramáticas de unificación PC-PARSE. Asimismo, se incluyen programas ejemplo, desarrollados expresamente para el curso, con los que verificar determinados aspectos prácticos del curso.

IGUALDAD DE GÉNERO

En coherencia con el valor asumido de la igualdad de género, todas las denominaciones que en esta Guía hacen referencia a órganos de gobierno unipersonales, de representación, o miembros de la comunidad universitaria y se efectúan en género masculino, cuando no se hayan sustituido por términos genéricos, se entenderán hechas indistintamente en género femenino o masculino, según el sexo del titular que los desempeñe.