

PROGRAMA DE ASIGNATURA

1. NOMBRE DE LA ASIGNATURA

Introducción al análisis automático de la gramática

2. NOMBRE DE LA ASIGNATURA EN INGLÉS

Introduction to the automatic analysis of grammar

3. TIPO DE CRÉDITOS DE LA ASIGNATURA

SCT/ UD/ OTROS/

4. NÚMERO DE CRÉDITOS

10

5. HORAS DE TRABAJO PRESENCIAL DEL CURSO

3

6. HORAS DE TRABAJO NO PRESENCIAL DEL CURSO

6

7. DESCRIPCIÓN DE LA ASIGNATURA

En este seminario, de carácter teórico-práctico, se abordan contenidos fundamentales de la morfología y la sintaxis del español, desde una perspectiva formal, con especial foco en la jerarquía Chomsky-Schützenberger (1963), y dentro del marco de la lingüística computacional. En relación con ello, se presentan, por un lado, diversas problemáticas referentes a los mecanismos de formación de palabras y la manera en que estas se combinan y establecen diversas funciones sintácticas. Por otro lado, se plantean métodos de formalización de dichas estructuras con vistas al diseño de algoritmos computacionales que permitan su identificación en textos de lenguaje natural y su generación automática. Para esto último, se recurre al programa NooJ (Silberztein, 2016) un software que cuenta con diversos recursos como diccionarios electrónicos y gramáticas informáticas. Se espera que, al finalizar el curso, los estudiantes adquieran conocimientos básicos para el desarrollo de recursos de análisis automáticos en el plano gramatical.

8. OBJETIVO GENERAL DE LA ASIGNATURA

Analizar fenómenos morfológicos y sintácticos del español desde una perspectiva formal con vista al modelamiento computacional para su reconocimiento en textos de lenguaje natural y generación automática.



OBJETIVOS ESPECÍFICOS DE LA ASIGNATURA

- Conocer los procesos de formación de palabras, sintagmas y oraciones.
- Conocer las metodologías de análisis automático basado en reglas.
- Desarrollar algoritmos computacionales básicos para el análisis automático para fenómenos gramaticales específicos.

9. RESULTADOS DE APRENDIZAJE ESPERADOS

- Analizar estructuras morfológicas y sintácticas del español desde una perspectiva formal, identificando los principios jerárquicos que rigen su organización interna.
- Diseñar algoritmos computacionales simples que permitan la generación y el reconocimiento automático de estructuras gramaticales específicas en textos de lenguaje natural.
- Elaborar descripciones gramaticales formalizadas mediante el uso de recursos informáticos (gramáticas computacionales, diccionarios electrónicos), integrando teoría lingüística y modelamiento computacional.

10. SABERES / CONTENIDOS

Unidad 1: Representación formal del conocimiento lingüístico

- Introducción a la lingüística computacional
- Lenguajes naturales y lenguajes formales
- La jerarquía Chomsky-Schützenberger
- El análisis automático de textos basado en reglas: aplicaciones

Unidad 2: Análisis automático de la morfología

- Aspectos de la morfología flexiva y derivativa del español
- Gramáticas morfológicas informáticas para generación y reconocimiento automático
- Construcción de diccionarios electrónicos

Unidad 3: Análisis automático de la sintaxis

- El nivel sintagmático: sintagmas léxicos y funcionales
- La oración simple: formalización de la estructura argumental del predicado
- Gramáticas sintácticas informáticas para generación y reconocimiento automático



11. METODOLOGÍA

Este curso se desarrolla a través de la discusión teórica y resolución de ejercicios de análisis lingüístico en el aula, sobre la base de lecturas específicas de la bibliografía del programa; para ello, es necesario que los estudiantes cuenten con computador y el programa de análisis computacional NooJ. Las sesiones se estructuran en tres partes. En primer lugar, el docente expone un tema, el cual es comentado y complementado por los otros participantes a la luz de la bibliografía del curso. En segundo lugar, se realiza una actividad práctica, individual o grupal, para ello, se les instalarán las herramientas de análisis automático presentadas en el curso. Por último, el profesor cierra la sesión con una síntesis de los temas tratados.

12. METODOLOGÍAS DE EVALUACIÓN

El curso contempla las siguientes evaluaciones para su aprobación:

- Un trabajo práctico correspondiente a las primeras dos unidades, individual o grupal. Dicha actividad implicará el 40% de la nota final.
- Un trabajo final escrito individual que contemple un proyecto de investigación en el ámbito de la gramática, dentro del marco del análisis computacional. Dicha actividad implicará el 60% de la nota final.

13. REQUISITOS DE APROBACIÓN

ASISTENCIA: 80%

NOTA DE APROBACIÓN MÍNIMA: 4.0

REQUISITOS PARA PRESENTACIÓN A EXAMEN: Tener bajo un 4.0

14. PALABRAS CLAVE

gramática, lingüística computacional, léxico, morfología, sintaxis

15. BIBLIOGRAFÍA OBLIGATORIA

- Bonino, R., & Rodrigo, A. (mimeo). NooJ: Tutorial en español.
- Bosque, I., & Demonte, V. (1999). Gramática Descriptiva de la Lengua Española, vol. 2: Las construcciones sintácticas fundamentales. Relaciones temporales, aspectuales y modales. Madrid: Espasa. ISBN: 84-239-7917-2
- Bosque, I., & Gutiérrez-Rexach, J. (2011). Fundamentos de sintaxis formal. Madrid: AKAL Editores. ISBN: 978-84-460-2227-5
- Chomsky, N. (1969). Syntactic structures. La Haya: Mouton. ISBN 978-3-11-021832-9
- Di Tullio, A. (2007). Manual de gramática del español. Buenos Aires: Waldhuter Ediciones.
- Fábregas, A. (2013). La morfología. El análisis de la palabra compleja. Madrid: Editorial Síntesis.
- Gallego, A. (2015). Perspectivas de sintaxis formal. Madrid: AKAL.

- Gross, G. (2014). *Manual de análisis lingüístico*. Barcelona: Editorial UOC.
- Gross, M. (1984). *Lexicon-Grammar and the syntactic analysis of French*. En *International Conference on Computational Linguistics (COLING)*. Stanford, California: Stanford University.
- Gross, M. (1994). *The lexicon-grammar of a language*. En *Encyclopedia of language*. Oxford: Pergamon Press.
- Koza, W., & Mare, M. (2023). *Computational modeling of a generative grammar for clitic order in River Plate Spanish*. *RESLA*.
- La Serna, N. (2004). *Un analizador sintáctico eficiente para gramáticas del español*. *RISI*, 1(1), 19-26.
- Mella Llopart, M. (1983). *Análisis morfológico automático del español*. *Procesamiento del lenguaje natural*, 1, 18-23.
- Moreno Sandoval, A. (1998). *Lingüística computacional*. Madrid: Editorial Síntesis.
- Moreno Sandoval, A. (2019). *Lenguas y computación*. Madrid: Editorial Síntesis.
- Periñán, C. (2012). *En defensa del procesamiento del lenguaje natural fundamentado en la lingüística teórica*. *Onomazein*, 26, 13-48.
- RAE. (2009). *Nueva gramática de la lengua española*. Madrid: Espasa.
- Rivera, R., Barrón, C., & Calvo, H. (2016). *Análisis y comparación de diferentes métodos de reconstrucción de árboles semánticos*. *Research in Computing Science*, 120, 81-88.
- Silberztein, M. (2003). *NooJ manual*. Disponible en: <http://nooj4nlp.org/downloads.html>
- Silberztein, M. (2016). *Formalizing natural languages: The NooJ approach*. Londres: ISTE.
- Solias Arís, T. (2015). *Métodos formales en lingüística*. Madrid: Editorial Síntesis.
- Varela, S. (2005). *Morfología léxica: la formación de palabras*. Madrid: Gredos.

16. BIBLIOGRAFÍA COMPLEMENTARIA

- Abney, S. (1992). *Parsing by chunks*. In R. Berwick, S. Abney, & C. Tenny (Eds.), *Principle Base parsing*. *Studies in Linguistics and Philosophy*, 44, 257-278.
- Beesley, K., & Karttunen, L. (2003). *Finite-state morphology*. Stanford: CSLI.
- Bès, G. (1999). *La phrase verbale noyau en français*. *Recherches sur le français parlé*, 15, 273-358.
- Bès, G., & Solana, Z. (2007). *Sintagma Verbal Núcleo flexionado en español*. *Infosur Revista*, 1, 15-26.
- Blanco, X. (2016). *A hierarchy of semantic labels for Spanish dictionaries*. *Communications in Computer and Information Science*, 667, 66-73.
- Bonino, R. (2015). *Una propuesta para el tratamiento de los enclíticos en NooJ*. *Infosur Revista*, 7, 31-40.
- Castel, V. (1994). *Paquets des clitiques de l'espagnol: calcul de la corrélation entre rôles sémantiques et propriétés morphologiques et implementations dans une grammaire catégorielle d'unification* (Doctoral dissertation). Université Blaise Pascal.
- Castel, V. (2007). *An OpenCCG Implementation of Perlmutter's Generalization on Spanish Clitics Word Order*. *Infosur Revista*, 1, 3-14.
- Chomsky, N. (1970). *Remarks on nominalizations*. In R. Jacobs & P. Rosenbaum (Eds.), *Readings in English transformational grammar* (pp. 184-221). Waltham: Ginn.
- De Miguel, E. (2009). *La Teoría del Lexicón Generativo*. In E. De Miguel (Ed.), *Panorama de la lexicología* (pp. 337-368). Madrid: Ariel.
- De Miguel, E. (2011). *En qué consiste ser verbo de apoyo*. In V. Escandell, M. Leonetti, & C. Sánchez (Eds.), *60 problemas de gramática* (pp. 139-146). Madrid: Akal.
- De Miguel, E., & Fernández Lagunilla, M. (2007). *Sobre la naturaleza léxica del aspecto composicional*. *Actas del VI Congreso de Lingüística General* (pp. 1767-1778). Madrid: Arco/Libros.
- Elia, A., Monteleone, M., & Marano, F. (2011). *From the concept of transformation in Harris and Chomsky to the Lexique-Grammaire of Maurice Gross*. In V. Kasevich, Y. Kleiner, & P. Sériot (Eds.), *History of linguistics* (pp. 76-82). Amsterdam: John Benjamins.
- Grimshaw, J. (1990). *Argument Structure*. Cambridge, MA: MIT Press.
- Harris, Z. (1970). *Papers in structural and transformational linguistics*. Dordrecht: Springer.

- Kocijan, K., & Librenjak, S. (2016). Recognizing verb-based Croatian idiomatic MWUs. *Communications in Computer and Information Science*, 667, 96-106.
- Lamiroy, B. (1991). *Léxico y gramática del español. Estructuras verbales de espacio y tiempo*. Barcelona: Editorial Anthropos.
- Langella, A. (2016). Paraphrases for the Italian communicative predications. *Communications in Computer and Information Science*, 811, 182-195.
- Le Pesant, D., & Mathieu-Colas, M. (1998). Introduction aux classes d'objets. *Langages*, 131, 6-33.
- Marantz, A. (1984). *On the Nature of Grammatical Relations*. Cambridge, MA: MIT Press.
- Messina, S., & Langella, A. (2015). Paraphrases $V \leftrightarrow N \leftrightarrow A$ in one class of Psychological Predicates. In J. Monti, M. Silberztein, M. Monteleone, & M. di Buono (Eds.), *Formalising Natural Languages with NooJ 2014* (pp. 140-149). Newcastle: Cambridge Scholars Publishing.
- Perlmutter, D. (1971). *Deep and Surface Structure Constraints in Syntax*. New York: Holt, Rinehart and Winston.
- Pineda, L., & Meza, I. (2005). The Spanish pronominal clitic system. *Sociedad Española para el Procesamiento del Lenguaje Natural*, 35, 1-37.
- Pustejovsky, J. (1995). *The generative lexicon*. Cambridge, MA: MIT Press.
- Tolone, E. (2009). Les tables du Lexique-Grammaire au format TAL. Ponencia presentada en MajecsTIC 2009, Avignon, Francia.
- Zeevat, H., Klein, E., & Calder, J. (1987). Unification categorical grammar. In N. Haddock, E. Klein, & G. Morrill (Eds.), *Edinburg working papers in cognitive science: Categorical grammar, unification grammar, and parsing* (Vol. 1, pp. 195-222). Edinburg: Centre for Cognitive Science, University of Edinburg.

17. RECURSOS WEB

ht tps://nooj.univ-fcomte.fr/

Contacto: walter.koza@uchile.cl