



Universidad Nacional del Comahue
Facultad de Lenguas
Ciclo lectivo 2024

SECRETARÍA DE POSGRADO

PROGRAMA

1. Título del seminario:

Seminario Optativo de Maestría - Seminario de Investigación Lingüística I o II
(*Lingüística Computacional*)

2. Carreras:

Maestría en Lingüística

3. Tipo de actividad:

Seminario

4. Fecha:

Desde el **27 de abril** al **18 de mayo** de 2024

5. Modalidad de dictado:

Presencialidad Virtual

6. Lugar de dictado:

Facultad de Lenguas, UNComahue

7. Días y Horarios:

Días: **sábados de 9:00 a 12:00 horas**

Fechas:

27 de abril

04 de mayo

11 de mayo

18 de mayo

1° de junio

8. Duración en horas:

50 horas - Las horas están distribuidas en 15 horas de cursado y 4 horas de consulta ambas de carácter presencial virtual. Las horas restantes de cursado serán de carácter asincrónico





(resolución de actividades, intercambio a través de la plataforma PEDCO, lectura bibliográfica, uso de foros, etc).

9. Disertantes

Dr. Walter Adrián Koza (CONICET - UNGS)

9.1. Curriculum vitae sintético de cada uno de los profesores participantes

Walter A. Koza es Profesor y Licenciado en Letras, Doctor en Humanidades y Artes con mención en Lingüística, egresado de la Universidad Nacional de Rosario (Argentina). Fue profesor adjunto del Instituto de Literatura y Ciencias del Lenguaje, de la Pontificia Universidad Católica de Valparaíso, Chile y, actualmente, es investigador adjunto del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Sus intereses se centran en la lingüística computacional, la gramática del español y las alteraciones del lenguaje, áreas en las que ha publicado diversos artículos.

10. Fundamentación

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.

11. Objetivos

Objetivo general

- 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:

- Conocer los procesos de formación de palabras, sintagmas y oraciones
- Conocer las metodologías del análisis automático basado en reglas
- Desarrollar algoritmos computacionales básicos para el análisis automático de aspectos específicos de la gramática del español

12. Contenidos y bibliografía

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 del español

- 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 del español

- 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

Unidad 1: Representación formal del conocimiento lingüístico

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*. Espasa. Madrid.

Bosque, I. & Gutiérrez-Rexach, J. (2011). *Fundamentos de sintaxis formal*. Madrid: AKAL Editores.

Chomsky, N. (1969). *Syntactic structures*. Mouton. La Haya.



- Di Tullio, A. (2007). *Manual de gramática del español*. Buenos Aires: Waldhuter Ediciones.
- Gallego, A. (2015). *Perspectivas de sintaxis formal*. Madrid: AKAL.
- Moreno Sandoval, A. (2019). *Lenguas y computación*. Madrid: Editorial Síntesis.
- Moreno Sandoval, A. (1998). *Lingüística computacional*. 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.
- Silberztein, M. (2016). *Formalizing natural languages : The NooJ approach*. Londres: ISTE.
- Silberztein, M. (2003). *NooJ manual*. [En línea] Disponible en: <http://nooj4nlp.org/downloads.html>
- Solias Arís, T. (2015). *Métodos formales en lingüística*. Madrid: Editorial Síntesis

Unidad 2: Análisis automático de la morfología del español

- 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*. Espasa. Madrid.
- Bosque, I. & Gutiérrez-Rexach, J. (2011). *Fundamentos de sintaxis formal*. Madrid: AKAL Editores.
- Di Tullio, A. (2007). *Manual de gramática del español*. Buenos Aires: Waldhuter Ediciones.
- Fábregas, A. (2013). *La morfología. El análisis de la palabra compleja*. Editorial Síntesis. Madrid.
- Mella Llopart, M. (1983). Análisis morfológico automático del español. *Procesamiento del lenguaje natural*, 1, 18-23.
- RAE (2009). *Nueva gramática de la lengua española*. Madrid: Espasa.
- Silberztein, M. (2016). *Formalizing natural languages: The NooJ approach*. Londres: ISTE.
- Silberztein, M. (2003). *NooJ manual*. [En línea] Disponible en: <http://nooj4nlp.org/downloads.html>
- Varela, S. (2005). *Morfología léxica: la formación de palabras*. Madrid: Gredos.



Unidad 3: Análisis automático de la sintaxis del español

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*. Espasa. Madrid.

Bosque, I. & Gutiérrez-Rexach, J. (2011). *Fundamentos de sintaxis formal*. Madrid: AKAL Editores.

Di Tullio, A. (2007). *Manual de gramática del español*. Buenos Aires: Waldhuter Ediciones.

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)*. California: Stanford University.

Gross, M. (1994). The lexicon-grammar of a language. *Encyclopedia of language*. Oxford: Pergamon Press.

Koza, W. & Mare, M. (en prensa). 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.

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. (2016). *Formalizing natural language: The NooJ approach*. Londres: ISTE.

Silberztein, M. (2003). *NooJ manual*. [En línea] Disponible en: <http://nooj4nlp.org/downloads.html>

13. Metodología de trabajo y/o Actividades de los asistentes

Este curso se desarrolla a través de la discusión y práctica de los programas de análisis lingüístico en el aula, sobre la base de lecturas, exposiciones orales y ejercicios prácticos; 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, se 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.

Descripción de las actividades y tareas

La asignatura contempla la elaboración de implementaciones computacionales básicas para el abordaje de problemas lingüísticos de índole gramatical. Los estudiantes deberán diseñar los recursos pertinentes (diccionarios, gramáticas, corpus, etcétera) y evaluar sus alcances y limitaciones. Para ello, es necesario que cuenten con computadora, en la primera sesión se darán las indicaciones para la instalación de la herramienta NooJ (Silberztein, 2016).

14. Evaluación final

Para la regularización del seminario, se deberá realizar un trabajo práctico (individual o grupal). Para su aprobación, además, un trabajo final que, también, se podrá hacer tanto individualmente como en grupo.

15. Observaciones (si corresponde) (*)

Durante el dictado del seminario, se pactará entre el docente y los estudiantes espacios para clases de consulta.

Walter Koza