

**Formulario de Aprobación Curso de Posgrado**

**Asignatura:**

Gramáticas Formales para el Lenguaje Natural

---

**Profesor de la asignatura <sup>1</sup>:**

(título, nombre, grado o cargo, Instituto o Institución)  
Dra.Ing. Aiala Rosá, Prof. Adjunta. Instituto de Computación (Responsable)  
Dra.Ing. Dina Wonsever, Prof.Titular. Instituto de Computación

**Profesor Responsable Local <sup>1</sup>:**

(título, nombre, grado, Instituto)

**Otros docentes de la Facultad:**

(título, nombre, grado, Instituto)

**Docentes fuera de Facultad:**

(título, nombre, cargo, Institución, país)

**Programa(s):** Maestría en Informática PEDECIBA y Doctorado en Informática PEDECIBA

**Instituto ó Unidad:** Instituto de Computación

**Departamento ó Area:** Grupo PLN

<sup>1</sup> Agregar CV si el curso se dicta por primera vez.

(Si el profesor de la asignatura no es docente de la Facultad se deberá designar un responsable local)

---

**Horas Presenciales: 50**

(se deberán discriminar las mismas en el ítem Metodología de enseñanza)

**Nº de Créditos: 10**

(de acuerdo a la definición de la UdelaR, un crédito equivale a 15 horas de dedicación del estudiante según se detalla en el ítem metodología de la enseñanza)

**Público objetivo y Cupos:**

Estudiantes de posgrado en Informática. Estudiantes de posgrado en Lingüística. No existe cupo.

---

**Objetivos:**

Introducción a los principales temas y aplicaciones en Procesamiento de Lenguaje Natural y al rol de las gramáticas formales en este contexto. Conocimiento de distintos enfoques y formalismos para la descripción sintáctico-semántica del lenguaje natural. Experimentación con herramientas y gramáticas existentes. Adaptación a otras lenguas o dominios particulares de recursos gramaticales. Conocimiento de medidas de evaluación, trabajo con conjuntos de prueba.

---

**Conocimientos previos exigidos:**

Lógica de predicados.  
Probabilidad y estadística.  
Gramáticas formales.  
Programación.

---

**Conocimientos previos recomendados:**

---

**Metodología de enseñanza:**

---

(comprende una descripción de la metodología de enseñanza y de las horas dedicadas por el estudiante a la asignatura, distribuidas en horas presenciales -de clase práctica, teórico, laboratorio, consulta, etc.- y no presenciales de trabajo personal del estudiante)

La modalidad del curso se basa fuertemente en el trabajo del estudiante durante todo el semestre. Los estudiantes deben estudiar los temas teóricos en base a diferentes materiales (libros, artículos, casos filmadas, resúmenes en formato presentación) que los docentes pondrán a su disposición. En las clases presenciales se discutirán conceptos teóricos, se analizarán ejemplos y se resolverán ejercicios. También se responderán consultas de los estudiantes.

En paralelo a las actividades mencionadas, los estudiantes trabajarán en tareas de laboratorio que les permitirán experimentar con herramientas y recursos vinculados a los diferentes formalismos presentados y al proceso de análisis sintáctico.

- Horas clase (teórico): 10
- Horas clase (práctico): 20
- Horas clase (laboratorio):
- Horas consulta: 17
- Horas evaluación: 3
  - Subtotal horas presenciales: 50
- Horas estudio: 30
- Horas resolución ejercicios/prácticos: 10
- Horas proyecto final/monografía: 60
  - Total de horas de dedicación del estudiante: 150

---

### Forma de evaluación:

El curso se evaluará a partir de:

- trabajos de laboratorio: 50%
- prueba individual escrita: 50%

Las dos instancias son obligatorias y se requerirá como mínimo el 60% del puntaje de la prueba individual escrita. El curso se aprueba obteniendo más del 60% de los puntos totales.

---

### Temario:

1. Introducción al área Procesamiento de Lenguaje Natural (PLN). Motivación. Aplicaciones más comunes.
2. Conceptos básicos de gramática. Nociones básicas de gramática del español.
3. Presentación de distintos formalismos gramaticales:
  - Gramáticas libres de contexto.
  - Gramáticas libres de contexto probabilísticas.
  - Gramáticas de unificación, HPSG.
  - Gramáticas de dependencias.
  - Gramáticas categoriales.
4. Análisis sintáctico. Métodos de parsing. Medidas de evaluación.
5. Corpus para el desarrollo de gramáticas, conjuntos de prueba.
6. Formalismos semánticos, representaciones AMR.

**Bibliografía:**

(título del libro-nombre del autor-editorial-ISBN-fecha de edición)

---

**Básica**

1. Di Tullio, Ángela; Marisa Malcuori (2012). Gramática del español para maestros y profesores del Uruguay. Montevideo: ANEP. ProLEE. ISBN 978-9974-688-79-7.
2. Jurafsky, Daniel, James H. Martin (2008). Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition (2<sup>nd</sup> edition). New Jersey: Prentice Hall. ISBN-10: 0131873210.
3. Melcuk, Igor (1988). Dependency Syntax: Theory and Practice. New York: State University of New York Press.
4. Sag, I. A., Wasow, T., Bender, M. E. (2003). "Syntactic Theory. A Formal Introduction". 2<sup>a</sup> edición. CSLI Publications. Stanford.
5. Steedman, M. y J. Baldridge (2011). Combinatory Categorical Grammar, en R. Borsley and K. Borjars (eds.) Non-Transformational Syntax, 181-224, Blackwell.
6. Banarescu, L., Bonial, C., Cai, S., Georgescu, M., Griffitt, K., Hermjakob, U., Knight, K., Koehn, P., Palmer, M., and Schneider, N. (2013). Abstract meaning representation for semantic, en Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse. Association for Computational Linguistics.

**Complementaria**

7. Carpenter. B. (1992). The Logic of Typed Feature Structures, B. Carpenter, Cambridge University Press, USA, 1992. ISBN 0521419328.
  8. Copestake, Ann (2002). Implementing Typed Feature Structure Grammars (University of Cambridge) Stanford, CA: CSLI Publications (CSLI lecture notes, number 110), ISBN 1-57586-261-1.
  9. Costa, S. y M. Malcuori (1990). Introducción a la Teoría del Gobierno y la Ligadura de Noam Chomsky. Serie "Avances de Investigación", Fac. de Humanidades y Ciencias. Montevideo.
  10. Hockenmaier, J. (2003). Data and Models for Statistical Parsing with Combinatory Categorical Grammar, PhD thesis, University of Edinburgh
  11. Hockenmaier, J. y M. Steedman. (2007). CCGBank: A Corpus of CCG Derivations and Dependency Structures Extracted from the Penn Treebank. Computational Linguistics. Volumen 33, Nro. 3.
  12. Nivre, J., Hall, J., Nilsson, J., Chanev, A., Eryigit, G., Kübler, S., Marinov, S. and Marsi, E. (2007). MaltParser: A language-independent system for data-driven dependency parsing. Natural Language Engineering, 13(2), 95-135
  13. Nivre, J. (2008). Algorithms for Deterministic Incremental Dependency Parsing. Computational Linguistics 34(4), 513-553
  14. Pollard, C. J. y Sag, I. A. (1987). "Information -Based Syntax and Semantics: Volume I, Fundamentals", Volumen 13 de CSLI Lecture Notes. Center for the Study of Language and Information. Stanford.
  15. Steedman, M. (1996). A Very Short Introduction to CCG, en línea: <http://www.inf.ed.ac.uk/teaching/courses/nlg/readings/ccgintro.pdf>
  16. Tesnière, Lucien (1959). Éléments de syntaxe structurale. Editions Klincksieck.
-