

Universidad de Puerto Rico
Recinto Universitario de Mayagüez
Facultad de Artes y Ciencias
Departamento de Estudios Hispánicos

PROPUESTA PARA LA
SECUENCIA CURRICULAR EN LINGÜÍSTICA COMPUTACIONAL

El Departamento de Estudios Hispánicos propone la creación de una Secuencia Curricular en Lingüística Computacional como parte de los ofrecimientos de su Programa de Bachillerato. Esta secuencia cumple con los requisitos académicos de una secuencia curricular de categoría IV según la Certificación Número 15-07 (2014-2015) del Senado Académico del Recinto Universitario de Mayagüez por ser una secuencia articulada e integrada de cursos que proveen una alternativa pertinente, innovadora e interdisciplinaria a la lingüística computacional. La secuencia constituye un valor añadido no solamente al bachillerato en Estudios Hispánicos, Matemáticas, Ciencias de Computación, Ingeniería de Computadoras e Ingeniería de Software, sino también a todos los programas de bachillerato centrados en el estudio de las ciencias cognitivas (lingüística, psicología, filosofía, inteligencia artificial), ciencias naturales (matemáticas, biología), ciencias sociales (sociología, antropología) y otros idiomas (inglés, francés, italiano, etc.).

I. Objetivos

Esta Secuencia Curricular en Lingüística Computacional tiene los siguientes objetivos:

1. Diversificar y enriquecer la preparación académica y profesional de los estudiantes del Programa de Estudios Hispánicos y de otros programas del Recinto Universitario de Mayagüez, exponiéndolos al estudio de las propiedades computacionales del lenguaje humano y de los modelos de conocimiento y procesamiento de lenguas naturales.
2. Facilitar la admisión de egresados del bachillerato a programas graduados de lingüística teórica, lingüística computacional, procesamiento de lenguas naturales e inteligencia artificial, que requieren una formación subgraduada inexistente hasta ahora en nuestro recinto y en los demás recintos de la Universidad de Puerto Rico.
3. Promover la inserción de egresados del bachillerato en profesiones y áreas académicas relacionadas con el estudio de la lingüística teórica y computacional, adquisición y procesamiento de lenguas naturales, programación, sistemas expertos e inteligencia artificial, brindándoles los conocimientos y herramientas indispensables para colaborar en ambientes multidisciplinarios junto a profesionales de las ciencias cognitivas, naturales y sociales, la educación y la ingeniería de computación y software.
4. Contribuir a la formación integral de estudiantes de ingeniería, ciencias naturales, ciencias sociales y humanidades, ofreciéndoles una perspectiva cognitiva a través de la lingüística en su formación académica y quehacer profesional.

Esta secuencia sirve de complemento ideal para la formación de estudiantes egresados de otras disciplinas en esta universidad que deseen ensayar y desarrollar modelos de adquisición y procesamiento lingüístico, así como estrategias del método científico en un estudio formal más allá de las fronteras de sus áreas inmediatas de especialidad.

II. Justificación

Esta secuencia curricular le ofrece al estudiantado una formación competitiva en las áreas esenciales de la teoría y práctica de la lingüística computacional y el procesamiento de lenguas naturales. El conocimiento de la teoría lingüística, particularmente la sintaxis y la semántica, capacita al estudiante para entender los principios universales y los parámetros de variación en las lenguas naturales, las propiedades y rasgos del lenguaje como facultad exclusiva de la especie humana y las gramáticas como las representaciones mentales de un sistema cognitivo computacional. El estudio de los fundamentos formales de la lingüística computacional le brinda al estudiante las herramientas lógico-matemáticas esenciales para el análisis y evaluación de modelos computacionales de aprendizaje y procesamiento lingüístico basados en sistemas deterministas y no-deterministas, simbólicos y probabilísticos y le permite familiarizarse con herramientas en línea para procesamiento de lenguas naturales (PLN), tales como corpus anotados, analizadores estructurales y redes y ontologías semánticas. La capacidad para programar en lenguajes de computación procedimentales y declarativos y manejar diferentes técnicas y formatos de representación, almacenamiento y recuperación de información es indispensable para el desarrollo de modelos computacionales de PLN.

Los sistemas de PLN tienen aplicaciones cada vez más necesarias (Moreno Sandoval 1998), tales como: (i) *traducción automática* a partir de textos escritos o de emisiones orales; (ii) *recuperación de información*, que requiere sistemas que realmente entiendan las condiciones de búsqueda y sean capaces de reconocer qué documentos son relevantes o no; (iii) *extracción de información*, que requiere reconocer la información relevante en una base de datos para trasladarla a formatos predeterminados (como tablas o gráficas); (iv) *interfaces hombre-máquina*, incluyendo agentes conversacionales (como *Siri* en iOS), en las que el usuario use una lengua natural en vez de una artificial o un menú restringido de opciones; (v) *correctores ortográficos*, que deben poseer al menos conocimiento de análisis morfológico y estructura silábica; (vi) *correctores sintácticos y de estilo*, que necesitan conocimiento sintáctico para detectar fallos de concordancia y oraciones “incompletas” o “incorrectas”, (vii) *enseñanza asistida computarizada de lenguas*, que deben tener capacidad de análisis sintáctico para plantear y corregir ejercicios de gramática y composición y (viii) *reconocimiento y síntesis de habla y texto* (para agentes conversacionales, sintetizadores de voz y transcritores como *Dragon*), que requieren conocimiento sintáctico para procesar aspectos prosódicos (como la entonación). Actualmente hay una gran demanda de sistemas de PLN, sobre todo para lenguas como el español, la segunda lengua con más hablantes nativos y la tercera más hablada del mundo (Ethnologue). Se estima que para empresas como Amazon, Apple, Facebook, Google, IBM e Intel, el ingreso anual de tecnologías y servicios de PLN ascenderá en el 2024 a 2.1 billones de dólares (Dataversity). Nuestro recinto, comprometido con el bienestar y desarrollo de nuestro país, cuenta con los recursos físicos e intelectuales necesarios para ofrecerles a nuestros estudiantes una oferta académica que les permita colocarse a la vanguardia en esta línea de investigación de tanto potencial y beneficio.

Esta secuencia fortalecerá significativamente la preparación académica y las competencias de los egresados del bachillerato en Estudios Hispánicos, ciencias de computación e ingeniería de computadoras y software para ser aceptados en programas graduados de lingüística teórica, lingüística computacional y ciencias cognitivas y para ser empleados como profesionales de la informática, el desarrollo de sistemas expertos y el procesamiento de lenguas naturales. De igual

manera, les provee a los egresados de programas en las humanidades, las ciencias sociales y otras especialidades la oportunidad de integrar en su labor profesional métodos y avances de disciplinas complementarias que abonan a su efectividad y a su capacidad para pensar de manera creativa e innovadora.

Entre las destrezas específicas que los estudiantes adquirirán mediante esta secuencia se incluyen:

1. Destrezas de análisis de datos lingüísticos mediante el uso de métodos formales para descomponer estructuras lingüísticas en sus elementos constituyentes, identificar patrones estructurales, establecer un modelo que permita la representación más adecuada de estas estructuras y formular reglas y generalizaciones con capacidad descriptiva y explicativa que den cuenta de los datos observados.
2. Destrezas de programación en lenguajes como Python, Prolog, Java u otros, que provean la experiencia necesaria para crear algoritmos eficientes, estructuras de datos y programas robustos y para utilizar las herramientas existentes de PLN.
2. Destrezas de investigación en lingüística teórica y computacional para diseñar sistemas computacionales de lenguaje, ya sean simbólicos o probabilísticos, que sirvan como modelos de adquisición o procesamiento lingüístico, así como para redactar informes y/o monografías científicas, usando el discurso de la disciplina, la terminología especializada y el estilo apropiado.

III. Cursos en la secuencia (18 créditos)

La Secuencia Curricular en Lingüística Computacional consta de 18 créditos. Puesto que la Universidad de Puerto Rico actualmente no ofrece un programa académico similar, una secuencia de 18 créditos sería homologable en otras universidades como una especialidad secundaria (“minor”). Esto sería conveniente para los estudiantes que quieran solicitar a un programa graduado de lingüística o de ciencias de computación, ya que los programas más competitivos dan preferencia a los candidatos que tengan una especialidad afín a estas disciplinas. En estos programas, una especialidad secundaria en lingüística o ciencias de computación requiere entre 15 a 18 créditos. La secuencia está diseñada tomando en cuenta los requisitos de otros programas análogos, el rigor de nuestros programas y la riqueza de nuestra oferta académica. De esta manera, los egresados de nuestro recinto, al momento de solicitar admisión a un programa graduado de lingüística computacional o teórica, no confrontarían la desventaja que les supondría no tener una secuencia curricular homologable.

En esta secuencia, cada estudiante deberá tomar un total de 9 créditos en cursos medulares, 6 créditos en requisitos de área y 3 créditos en una electiva recomendada para la secuencia.

CURSOS MEDULARES (9 CRÉDITOS):

- **LING 5030 Introducción a la Sintaxis Generativa** (3 créditos), que consiste en el estudio y la representación de las estructuras sintácticas de lenguas naturales (con énfasis en el español) y expone al estudiante al análisis de constituyentes y operaciones mediante la aplicación de las teorías sintácticas generativas más recientes.
- **LING 5060 Semántica composicional** (3 créditos), ya que el estudio de la composición del significado lingüístico es fundamental para elaborar una teoría sobre la competencia lingüística de los hablantes que pueda representarse mediante un modelo computacional.

• **LING 5080 Lingüística computacional** (3 créditos) complementa los cursos de sintaxis y semántica con el estudio de las propiedades computacionales del lenguaje humano y de los modelos de procesamiento de lenguas naturales mediante el análisis y las aplicaciones de modelos computacionales de aprendizaje y procesamiento lingüístico.

REQUISITOS DE ÁREA (6 CRÉDITOS):

• **Fundamentos formales** (3 créditos). Provee las herramientas básicas de lógica simbólica, matemática discreta y teoría de lenguajes formales necesarias tanto para la formulación de la teoría y las representaciones sintácticas y semánticas como para las ciencias de computación y programación. Este requisito puede satisfacerse con cualquiera de los cursos **LING 5090 Fundamentos formales de la teoría lingüística**, **CIIC 3075** o **ICOM 4075 Fundamentos de computación**, que enfatizan contenidos similares para estos propósitos.

• **Programación y estructura de datos** (3 créditos), Se presentan los mecanismos para la representación, el almacenamiento, la distribución, la recolección y el ordenamiento de datos en diversos formatos, imprescindibles en el diseño de un corpus etiquetado, un lexicón estructurado o un analizador sintáctico. Este requisito puede satisfacerse con cualquiera de los cursos **CIIC 4020/COMP 3075/ICOM 4035 Estructura de datos**.

ELECTIVA RECOMENDADA (3 CRÉDITOS):

El estudiante tomará un curso (3 créditos) como electiva recomendada de entre los siguientes:

- LING 4015 Seminario de Lingüística**
- LING 5040 Introducción a la fonología generativa**
- LING 5050 Teoría morfológica**
- LING 5120 Psicolingüística**
- CIIC 5015/COMP 5015/ICOM 5015 Inteligencia Artificial**
- CIIC 5045/COMP 5045 Lenguajes formales y autómatas**

A esta lista podrían añadirse otros cursos de LING o CIIC/COMP/ICOM que, previa aprobación, se estime pertinente por su vinculación con la lingüística computacional.

La tabla a continuación ilustra el programa de cursos en esta Secuencia:

CURSOS	CRÉDITOS
1er año de la secuencia curricular (correspondiente al 2do año académico del estudiante)	
LING 5030 Introducción a la Sintaxis Generativa	3
LING 5090 Fund. formales de la ling. o CIIC 3075/ICOM 4075 Fund. de comput.	3
2do año de la secuencia curricular (correspondiente al 3er año académico del estudiante)	
LING 5060 Semántica composicional	3
CIIC 4020/COMP 3075/ICOM 4035 Estructura de datos	3
3er año de la secuencia curricular (correspondiente al 4to año académico del estudiante)	
LING 5080 Lingüística computacional	3
LING 4015 o LING 5040 o LING 5050 o LING 5120 o CIIC 5015/COMP 5015/ICOM 5015 o CIIC 5045/COMP 5045	3

Todos los cursos de esta Secuencia Curricular en Lingüística Computacional son parte de la oferta académica regular del RUM.

IV. Requisitos mínimos para que el estudiante cualifique para declarar su intención de incluir la secuencia en sus planes de estudio

Podrá solicitar admisión a la Secuencia Curricular en Lingüística Computacional cualquier estudiante regular o de mejoramiento profesional del Recinto Universitario de Mayagüez. Al momento de declarar su intención de incluir la Secuencia Curricular en Lingüística Computacional dentro de sus planes de estudio, el estudiante debe haber completado un mínimo de 24 créditos con promedio de 2.50 o más en su programa de estudios y haber aprobado como prerrequisitos LING 4010 y (CIIC 3011 o COMP 3010 o INGE 3016) con una nota mínima de C en cada curso. El estudiante es responsable de cumplimentar y someter la solicitud de ingreso a la secuencia al Coordinador de la Secuencia Curricular en Lingüística Computacional para evaluar los méritos de la solicitud. En el caso de que un estudiante solicite readmisión al RUM, esta tiene que estar en conformidad con la fecha establecida por la Junta Administrativa. En los casos de traslado interno o externo, el estudiante tiene que cumplir con los requisitos establecidos para esos casos.

V. Requisitos de aprobación y residencia para completar la Secuencia satisfactoriamente

Para completar la Secuencia Curricular en Lingüística Computacional y obtener un certificado y anotación en el expediente, el estudiante habrá aprobado todos y cada uno de los cursos de la secuencia con calificación de C o más, cumpliendo con el requisito de residencia y habiendo obtenido un promedio de 2.50 o más en los cursos de la secuencia. Para cumplir el requisito de residencia, todos los cursos LING serán tomados y aprobados en el Departamento de Estudios Hispánicos del RUM. Una vez el estudiante haya cumplido satisfactoriamente con todos los requisitos de esta secuencia al momento de graduación, se hará constar mediante una anotación especial en el expediente del estudiante que incluirá el título de la secuencia y los cursos aprobados. La anotación de registraduría en la transcripción del estudiante leerá: "Successfully completed all requirements for the Curricular Sequence in Computational Linguistics."

VI. Plan de reclutamiento

Se diseñará e implementará un plan de reclutamiento para que los estudiantes de los programas de Estudios Hispánicos, Matemáticas, Ciencias de Computación, Ingeniería de Computadoras, Ingeniería de Software y otros programas conozcan la Secuencia Curricular en Lingüística Computacional temprano en su carrera universitaria y puedan solicitar admisión a partir de su segundo año de estudio. Este plan de reclutamiento comprenderá estrategias como diseminación a través de páginas de internet institucionales, presentaciones a estudiantes de nuevo ingreso y la distribución de hojas informativas entre el personal de orientación académica del Recinto.

VII. Recursos físicos

Como la Secuencia Curricular en Lingüística Computacional cuenta con el apoyo académico y administrativo del Departamento de Estudios Hispánicos, los estudiantes admitidos a la secuencia tienen acceso a las facilidades adscritas a este departamento, tales como el Laboratorio de computadoras, el Centro de Investigaciones Lingüísticas del Caribe (CILC). Como estudiantes de UPR-M, tienen también acceso a la Biblioteca General, al Centro de Redacción y a los demás centros de cómputos del Recinto.

VIII. Facultad del programa de Lingüística y Facultad adjunta

Nombre y rango	Preparación	Especialidad	Áreas de Investigación
DEPARTAMENTO DE ESTUDIOS HISPÁNICOS			
Hilton Alers-Valentín, <i>Catedrático</i>	Ph.D., University of Massachusetts at Amherst, 2000	Teoría fonológica, Sintaxis generativa, Semántica formal	Teoría de la Optimalidad; sintaxis y semántica de la negación y la cuantificación; analizadores simbólicos
Melvin González, <i>Catedrático Asociado</i>	Ph.D., Ohio State University, 2010	Sintaxis, Semántica, Pragmática, Español Caribeño	Sintaxis y semántica de cláusulas no-verbales y sintagmas nominales
Doris Martínez, <i>Catedrática</i>	Ph.D., Universidad Autónoma de Madrid, 1998	Pragmática, Análisis del discurso	Discurso periodístico y cibernético; conectores en el discurso oral
Alexandra Morales, <i>Catedrática Auxiliar</i>	Ph.D., University of Illinois at Urbana-Champaign, 2014	Adquisición de L1 y L2, Psicolingüística	Marca de objeto en el español caribeño; procesamiento lingüístico infantil
FACULTAD AFILIADA			
Nayda G. Santiago, <i>Catedrática (INEL)</i>	Ph.D., Michigan State University, 2003	Arquitectura y organización computacional	Sistemas de alto rendimiento computacional; procesamiento paralelo
J. Fernando Vega-Riveros, <i>Catedrático (ICOM)</i>	Ph.D., Syracuse University, 1989	Inteligencia Artificial, Procesamiento de Lenguas Naturales	Publicación digital; sistemas expertos; redes neurales, redes semánticas y de conocimiento
Bienvenido Vélez, <i>Catedrático Asociado (CIIC)</i>	Ph.D., Massachusetts Institute of Technology, 1999	Compiladores, estructura de lenguajes de programación	Bioinformática, calidad de software, hallazgo y recuperación de información

IX. Plan de Avalúo

Para el avalúo de la secuencia se recopilará la siguiente información:

1. número de estudiantes matriculados en la secuencia por año académico.
2. progreso académico de los estudiantes en la secuencia por año académico.
3. número de estudiantes que completan la secuencia por año académico y el tiempo que les toma para finalizar la misma.
4. satisfacción de los estudiantes en la secuencia mediante un cuestionario al finalizar la misma.
5. número de estudiantes que luego de completar la secuencia curricular continúen estudios graduados en áreas afines
6. notas obtenidas por los estudiantes en los cursos de la secuencia.
7. departamentos de procedencia de los estudiantes.
8. cursos electivos dentro de la secuencia más frecuentemente seleccionados.

El análisis se hará de acuerdo con las siguientes estrategias de avalúo.

<i>Objetivo operacional</i>	<i>Instrumento</i>	<i>Cómo medir el éxito</i>	<i>Persona(s) responsable(s) para su implementación</i>	<i>Itinerario de avalúo</i>
Generar una demanda estable para la secuencia	Registro de los estudiantes que han sido admitidos a la secuencia curricular	10 estudiantes admitidos a la secuencia durante los tres primeros años de implantación con tendencia de aumento en los años subsiguientes	Director, Director Asociado del Departamento, Consejero/a académico, Coordinador de la Secuencia	Anual
Tasa alta de retención	Proporción de estudiantes que completan la secuencia en tres años	Al menos 65% de los estudiantes que entran a la secuencia la habrá completado al graduarse de bachillerato	Director, Director Asociado, Consejero Académico del Departamento, Coordinador de la Secuencia	Anual
Impacto sobre egresados	Cuestionario de Satisfacción para estudiantes que completen la secuencia	Habrà una tendencia de aumento en los egresados que recomendarían a otro estudiante el seguir la secuencia curricular	Consejero Académico, Coordinador de la Secuencia	Anual
	Registro de egresados que continúen estudios graduados en áreas afines	Al menos 50% de los egresados prosiguen estudios graduados en la disciplina	Consejero Académico, Coordinador de la Secuencia	Bienal
	Cuestionario de efectividad	Al menos 65% de los egresados consideran que la secuencia ha sido efectiva como instrumento de mejoramiento profesional en su disciplina	Consejero Académico, Coordinador de la Secuencia	Bienal

X. Administración de la Secuencia curricular

La secuencia contará con un coordinador adscrito al Departamento de Estudios Hispánicos. El coordinador será seleccionado de entre los profesores que ofrezcan cursos y realicen investigaciones en la especialidad de la secuencia. El coordinador(a) ofrecerá orientación académica a los estudiantes que participen y llevará un registro de cumplimiento de los requisitos de la secuencia y del plan de avalúo.

ANEJO

I. Descripción en español

La Secuencia Curricular en Lingüística Computacional del Departamento de Estudios Hispánicos le ofrece al estudiantado del Recinto Universitario de Mayagüez una formación competitiva en las áreas esenciales de la teoría y práctica de la lingüística computacional y el procesamiento de lenguas naturales. El conocimiento de la teoría lingüística, particularmente la sintaxis y la semántica, capacita al estudiante para entender los principios universales y los parámetros de variación en las lenguas naturales, las propiedades y rasgos del lenguaje como facultad exclusiva de la especie humana y las gramáticas como las representaciones mentales de un sistema cognitivo computacional. El estudio de los fundamentos formales de la lingüística computacional le brinda al estudiante las herramientas lógico-matemáticas esenciales para el análisis y evaluación de modelos computacionales de aprendizaje y procesamiento lingüístico basados en sistemas deterministas y no-deterministas, simbólicos y probabilísticos y le permite familiarizarse con herramientas en línea para procesamiento de lenguas naturales (PLN), tales como corpus anotados, analizadores sintácticos y redes y ontologías semánticas. La capacidad para programar en lenguajes de computación declarativos y procedimentales y manejar diferentes técnicas y formatos de representación, almacenamiento y recuperación de información es indispensable para el desarrollo de modelos computacionales de PLN.

II. Descripción en inglés

The Curricular Sequence in Computational Linguistics at the Department of Hispanic Studies offers the students of the University of Puerto Rico at Mayagüez a competitive preparation in essential theoretical and practical areas of computational linguistics and natural language processing. The knowledge of linguistic theory, particularly syntax and semantics, enables students to understand the universal principles and variation parameters in natural languages, language properties and features as a human species specific faculty and grammars as mental representations of a computational cognitive system. The study of formal foundations of computational linguistics gives students the essential logical and mathematical tools to analyze and evaluate computational models of language processing and learning in both deterministic and non-deterministic, symbolical and probabilistic systems, and allows them to get acquainted with on-line natural language processing (NLP) tools, such as annotated corpora, syntactic parsers and semantic webs and ontologies. The ability to programming in declarative and procedural computer languages and handle different data representation, storage, and retrieval formats and techniques is indispensable to the development of NLP computational models.

III. Perfil del estudiante y del egresado

Esta Secuencia curricular está dirigida a aquellos estudiantes de Estudios Hispánicos, ciencias de computación e ingeniería de computadoras y software o de cualquier otra disciplina humanística, científica o tecnológica que estén interesados en el estudio de la lingüística computacional y el procesamiento de lenguas naturales. Se espera que al terminar los requisitos de la secuencia, los estudiantes participantes tengan las destrezas necesarias para enfrentarse a problemas relacionados con representaciones, estructuras y modelos de

procesamiento o implementación del lenguaje. Así mismo, se espera que puedan continuar estudios de postgrado en lingüística teórica, lingüística computacional y ciencias cognitivas o trabajar en ambientes académicos y/o profesionales que requieran colaboraciones interdisciplinarias junto a científicos naturales y sociales y peritos en computación, informática, sistemas expertos y procesamiento de lenguas naturales.

VI. Descripciones de los cursos de la especialidad

LING 4010 El lenguaje en la mente humana: una introducción a la lingüística (3 créditos)

Introducción a la lingüística como ciencia cognitiva mediante el estudio del lenguaje humano dentro del contexto más amplio de otras facultades biológicas y cognitivas. Aplicación del método científico en la construcción de una teoría lingüística con adecuación explicativa. Caracterización de los niveles de representación de estructuras lingüísticas, de las propiedades distintivas de la facultad de lenguaje y de las diferencias entre el lenguaje humano y otros sistemas naturales o artificiales. Exploración integrada de asuntos empíricos del análisis lingüístico, cuestiones filosóficas en el estudio del lenguaje y problemas específicos en el desarrollo de la teoría lingüística. Discusión de las implicaciones de los hallazgos en biolingüística cognitiva en algunos debates de la filosofía, la psicología, la biología y las ciencias de computación.

LING 4015 Seminario de Lingüística (3 créditos)

Exploración de un área de estudio en la lingüística actual. Lectura y discusión de fuentes primarias. Delimitación de un tema y diseño de un proyecto de investigación original. Se requiere la redacción y presentación del trabajo de investigación.

LING 5030 Introducción a la Sintaxis Generativa (3 créditos)

Estudio de las estructuras sintácticas en las lenguas naturales, con atención particular al español. Descripción y clasificación de rasgos, categorías, funciones y operaciones sintácticas. Representación de cláusulas subordinadas, clíticos, la negación y la estructura informacional de la oración. Análisis de constituyentes sintagmáticos y oracionales mediante la aplicación de los modelos teóricos más recientes en la sintaxis generativa.

LING 5040 Introducción a la Fonología Generativa (3 créditos)

Estudio de las estructuras fonológicas en las lenguas naturales, con atención particular al español. Descripción y clasificación de los sonidos articulados y rasgos distintivos de los segmentos y suprasegmentos. Representación de reglas fonológicas, geometrías de rasgos, entonación y acento métrico. Análisis de patrones y procesos fonológicos y estructuras prosódicas mediante la aplicación de las teorías fonológicas generativas más recientes.

LING 5050 Teoría Morfológica (3 créditos)

Representación de las estructuras, procesos y operaciones morfológicas en las lenguas naturales a la luz de los modelos teóricos propuestos en la morfología generativa. Estudio de la naturaleza del lexicón, la morfología como módulo autónomo de la gramática y la interfaz de la morfología con la fonología y la sintaxis. Revisión de las teorías de Morfología Léxica, Morfología Prosódica y Optimalidad. Aplicación de la teoría lingüística al análisis de datos en la morfología de lenguas naturales.

LING 5060 Semántica Composicional (3 créditos)

Introducción al estudio del significado lingüístico y su relación con la estructura sintáctica de acuerdo con los principios de la semántica composicional. Aplicación al análisis semántico de métodos formales y herramientas básicas, como la teoría de conjuntos, la lógica proposicional y la teoría de modelos. Exploración de los tipos y relaciones de significado, la predicación, la cuantificación, la modificación, las relaciones temporales, los contextos modales y la teoría de mundos posibles.

LING 5080 Lingüística computacional (3 créditos)

Introducción al estudio de las propiedades computacionales del lenguaje humano y de los modelos de procesamiento de lenguas naturales. Análisis y evaluación de sistemas deterministas y no-deterministas de modelos computacionales de aprendizaje y procesamiento lingüístico: autómatas y transductores, sistemas de reglas formales, sistemas lógico-matemáticos y modelos probabilísticos. Representación de estructuras fonológicas, morfológicas, sintácticas y semánticas mediante analizadores basados en estos modelos computacionales. Revisión de herramientas en línea como corpus anotados, procesadores estructurales y redes semánticas. Discusión de aplicaciones de modelos computacionales en tecnologías de procesamiento lingüístico, tales como síntesis y reconocimiento de habla, comprensión de lenguas naturales, correctores ortográficos y gramaticales, traducción automática, motores de búsqueda, desambiguación léxica, extracción de información, respuesta automática a preguntas y agentes conversacionales.

LING 5090 Fundamentos formales de la teoría lingüística (3 créditos)

Estudio de los fundamentos lógicos y matemáticos necesarios para formular la teoría lingüística y describir formalmente propiedades de lenguajes. Introducción a herramientas formales y conceptos básicos de teoría de conjuntos, relaciones y funciones; infinitos; cálculo proposicional y lógica de predicados; teoría de modelos; álgebras, reticulados, tipos de gramáticas, lenguajes formales y autómatas. Aplicación de métodos formales al análisis de la sintaxis y semántica de cuantificadores y propiedades de lenguajes naturales y formales.

LING 5120 Psicolingüística (3 créditos)

Introducción al estudio de los procesos y representaciones mentales involucrados en la implementación del lenguaje, que incluye la comprensión, producción y almacenamiento de información lingüística, tanto hablada como escrita. Revisión de los modelos de procesamiento estructural a nivel oracional, discursivo y conversacional. Exploración de la realidad psicológica de las representaciones lingüísticas, errores comunes de habla, resolución de ambigüedad, complejidad lingüística, el uso de información prosódica, léxica, sintáctica, semántica, pragmática y contextual en la comprensión lingüística y la relación entre los recursos computacionales disponibles en la memoria operativa y el mecanismo de procesamiento lingüístico. Discusión de las aportaciones de la psicología, las ciencias de computación y la inteligencia artificial en el diseño de modelos de procesamiento de lenguas naturales, así como las implicaciones de los hallazgos en psicolingüística para algunas de las controversias en estas disciplinas.

COMP 3075/CIIC 4020/ICOM 4035 Estructura de datos (3 créditos)

Estructuras de datos en lenguajes de programación: representación de información en forma de datos; listas de forma lineal, ortogonal, en sucesión y en arreglo; estructuras tipo árbol; técnicas para el almacenamiento, la distribución, la recolección y el ordenamiento de datos.

CIIC 3075 Fundamentos de computación (3 créditos)

Estructuras discretas en ciencia de computación e ingeniería con énfasis en destrezas de solución de problemas y algoritmos. Los temas incluyen: teoría de conjuntos, lógica y técnicas de demostración, teoría de grafos, computabilidad y probabilidad discreta aplicada a problemas de computación.

ICOM 4075 Fundamentos de computación (3 créditos)

Discusión de fundamentos matemáticos utilizados, comúnmente en ciencia e ingeniería con énfasis en destrezas de solución de problemas, algoritmos y modelos computacionales. Incluye tópicos tales como la relación entre datos y conjuntos, técnicas de demostración, operadores y funciones, lógica y circuitos elementales, grafos y organización de procesos computacionales, elementos de probabilidad discreta y eventos aleatorios según ocurren en computación. Los fundamentos matemáticos discutidos serán ilustrados con ejemplos extraídos de la disciplina de la computación.