



# UNIVERSITÀ DI PISA

---

## LINGUISTICA APPLICATA

**ALESSANDRO LENCI**

Anno accademico 2023/24  
CdS LINGUISTICA E TRADUZIONE  
Codice 1080L  
CFU 9

Moduli	Settore/i	Tipo	Ore	Docente/i
LINGUISTICA APPLICATA	L-LIN/01	LEZIONI	54	ALESSANDRO LENCI

### Obiettivi di apprendimento

#### *Conoscenze*

Lo studente acquisirà conoscenze in merito ai principali approcci alla rappresentazione semantica, con particolare riguardo alla semantica distribuzionale e alla sua applicazione per la modellazione linguistica e cognitiva. Lo studente verrà inoltre introdotto all'analisi statistica dei dati linguistici con l'ambiente R

#### *Modalità di verifica delle conoscenze*

L'accertamento delle conoscenze verrà svolto attraverso:

- prova orale, in cui verrà accertata la conoscenza dei testi e dei temi presentati a lezione
- prova scritta con la risoluzione di esercizi di analisi statistica con R

#### *Capacità*

Lo studente conoscerà i più recenti metodi e modelli di semantica e imparerà a usare l'ambiente R per l'analisi statistica dei dati linguistici

#### *Modalità di verifica delle capacità*

Verranno organizzate sessioni di laboratorio in cui saranno svolti esercizi per comprendere l'uso di software per la costruzione di spazi semantici distribuzionali e l'uso dell'ambiente di analisi statistica R

#### *Comportamenti*

Lo studente potrà acquisire e/o sviluppare sensibilità alle problematiche relative allo studio della semantica, con particolare riferimento ai modelli distribuzionali, e all'analisi statistica dei dati linguistici

Saranno acquisite accuratezza e precisione nell'analisi semantica e statistica dei dati linguistici

#### *Modalità di verifica dei comportamenti*

Durante le lezioni e le sessioni di laboratorio saranno valutati il grado di accuratezza e precisione delle attività svolte

#### *Prerequisiti (conoscenze iniziali)*

Nozioni di base di linguistica generale (morfologia, sintassi e semantica)

#### *Indicazioni metodologiche*

Lezioni: frontali

Modalità di apprendimento:

- frequenza delle lezioni
- studio individuale
- attività di laboratorio

Metodi di insegnamento:

- lezioni
- laboratorio

Lingua del corso: italiano



## UNIVERSITÀ DI PISA

---

### Programma (contenuti dell'insegnamento)

#### Modelli simbolici e distribuzionali della rappresentazione semantica

Il corso presenterà i principali modelli linguistici, computazionali e cognitivi per la rappresentazione semantica, con particolare attenzione ai modelli vettoriali in semantica distribuzionale:

- rappresentazioni simboliche del significato: semantica vero-condizionale, Conceptual Semantics
- problemi e limiti delle rappresentazioni simboliche
- fondamenti di semantica distribuzionale
- metodi per la costruzione di spazi semantici distribuzionali: matrici di co-occorrenza, modelli neurali (word embeddings), ecc.
- composizionalità semantica e modelli distribuzionali
- semantica distribuzionale e semantica formale
- applicazione dei modelli distribuzionali nelle (neuro)scienze cognitive;

Le lezioni teoriche saranno accompagnate da esercitazioni di laboratorio in cui verranno presentati semplici software per la costruzione di spazi semantici distribuzionali.

#### Esercitazioni:

Introduzione all'analisi statistica dei dati linguistici con R (<http://www.r-project.org/>):

- strutture dati
- statistica descrittiva (frequenze, tabelle di contingenza, medie, grafici, ecc.)
- test statistici di base (chi-2, t, ecc.)
- correlazione e regressione lineare

### Bibliografia e materiale didattico

Testi la cui conoscenza è oggetto di verifica all'esame:

Programma da 9 CFU:

- Pustejovsky J. & Batiukova O. (2019), *The Lexicon*, Cambridge, Cambridge University Press (ch. 4, 5, 6, 11)
- Lenci A. & Sahlgren M. (2023), *Distributional Semantics*, Cambridge University Press (ch. 1, 2, 6, 8.7, 8.8, 9)
- Yee, E., Jones, M. N., & McRae, K. (2018). "Semantic Memory". In J. T. Wixted & S. L. Thompson-Schill (Eds.), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience (4th Edition, Volume 3: Language and Thought)* (pp. 319–356). New York, NY: Wiley Blackwell.

Programma da 6 CFU:

- Pustejovsky J. & Batiukova O. (2019), *The Lexicon*, Cambridge, Cambridge University Press (ch. 4,5,11)
- Lenci A. & Sahlgren M. (2023), *Distributional Semantics*, Cambridge University Press (ch. 2, 6, 8.7, 9)

Testo di consultazione per le esercitazioni:

- Gries, S. Th. (2009), *Statistics for Linguistics with R*, Berlin, De Gruyter

### Indicazioni per non frequentanti

Testi la cui conoscenza è oggetto di verifica all'esame:

Programma da 9 CFU:

- Pustejovsky J. & Batiukova O. (2019), *The Lexicon*, Cambridge, Cambridge University Press (ch. 4, 5, 6, 11)
- Lenci A. & Sahlgren M. (2023), *Distributional Semantics*, Cambridge University Press (ch. 1, 2, 6, 8.7, 8.8, 9)
- Yee, E., Jones, M. N., & McRae, K. (2018). "Semantic Memory". In J. T. Wixted & S. L. Thompson-Schill (Eds.), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience (4th Edition, Volume 3: Language and Thought)* (pp. 319–356). New York, NY: Wiley Blackwell.

Programma da 6 CFU:

- Pustejovsky J. & Batiukova O. (2019), *The Lexicon*, Cambridge, Cambridge University Press (ch. 4,5,11)
- Lenci A. & Sahlgren M. (2023), *Distributional Semantics*, Cambridge University Press (ch. 2, 6, 8.7, 9)

Gli studenti non frequentanti sostituiscono la prova scritta su R con il seguente testo da portare all'esame orale:

- Pustejovsky J. & Batiukova O. (2019), *The Lexicon*, Cambridge, Cambridge University Press (cap. 8,9,10)

### Modalità d'esame

- Prova orale nella quale lo studente dovrà dimostrare la conoscenza dei temi svolti a lezione e dei testi di esame;



## UNIVERSITÀ DI PISA

---

- Prova scritta con la risoluzione di esercizi di analisi statistica con R

### Note

Commissione di esame:

- Prof. Alessandro Lenci
- Dott. Alessandro Bondielli
- Prof. Pietro Dini

Inizio lezioni: **26 febbraio 2024**

*Ultimo aggiornamento 18/02/2024 07:45*