Experimental and Computational Methods in Linguistic Research

Spring 2025

Instructor: Sanghee Kim

Week 1

Linguistic research?

- Linguistics?

- Research in Linguistics?

Some interesting questions

- When we hear the word *cat*, do we only think of this cute little furball?

Some interesting questions

- What happens in the brain when we listen to sentences like “I drank my coffee with sugar and cabbage,” or “All of the students was in the classroom”?

Some interesting questions

- The key to the cabinets are very rusty!
- The key to the cabinets is very rusty!

Case studies

- Above all the three case studies we will cover this quarter.

Subfields of linguistics

Phonetics & **phonology**: language forms
(sounds in spoken languages; signs in signed languages)

Semantics & **pragmatics**: linguistic meaning and use

Syntax & **morphology**: how forms combine to create more complex forms and meanings)

Psycholinguistics/Neurolinguistics: language, cognition, and the brain

Computational Linguistics: language and machines

Sociolinguistics: language and society

.. And more!

Psycholinguistics/Neurolinguistics

How is language implemented as a psychological process?

What constraints does the mind place on grammar?

What cognitive mechanisms underlie linguistic representation and processing?

Domains of questions in psycholinguistics

- **Comprehension**: decoding and understanding sounds, words, sentences, . . .
- **Production**: planning, constructing, and producing sounds, words, sentences, . . .
- **Acquisition**: how different aspects of language are acquired/learned

Four aspects of comprehension

How sounds are categorized

How words are accessed in the Lexicon

How syntactic structure is processed

How different types of meanings are processed

What methods can we use?

- **Behavioral methods**
  - Forced-choice categorization
  - Lexical decision with priming
  - Masked self-paced reading
  - Eye-tracking

What methods can we use?

- **Methods using neural measures**
  - Functional magnetic resonance imaging (fMRI), Positron emission tomography (PET)
  - Electroencephalography (EEG), Magnetoencephalography (MEG)
  - Learning from patients with brain damage

What methods can we use?

- **Computational methods**
  - Frequency information (text corpus data)
  - “Surprisal”

Conducting an experiment

- Research question
- Subjects
- Materials and design
- Procedure
- Results
- Analysis
- Report

What this course is about

- Research question
- Subjects
- **Materials and design ← Manual & Computational tools**
- Procedure
- Results
- Analysis
- Report

What this course is about

- Research question
- Subjects
- Materials and Design
- **Procedure** **← PCIbex; Lexical decision task & Self-paced reading**
- Results
- Analysis
- Report

What this course is about

- Research question
- Subjects
- Materials and Design
- Procedure
- **Results ← R; Data plotting (bar graph & line graph)**
- Analysis
- Report

What this course is about

- Research question
- Subjects
- Materials and Design
- Procedure
- **Results ← R; Data plotting (bar graph & line graph)**
- Analysis
- Report

What this course is about

- Research question
- **Subjects ← PCIbex; Replication study (mini-data collection)**
- Materials and Design
- Procedure
- Results
- Analysis
- Report

What this course is about

- **Research question**
- Subjects
- Materials and Design
- Procedure
- Results
- Analysis
- **Report**

What this course is about

- **Research question**
- Subjects
- Materials and Design
- Procedure
- Results
- Analysis
- **Report**

What this course is NOT about

- Research question
- Subjects
- Materials and Design
- Procedure
- Results
- Analysis ← We won’t have enough time to cover statistical analysis
- Report

What this course is NOT about

- Research question
- Subjects
- Materials and Design
- Procedure
- Results
- Analysis ← But we'll discuss the high-level concepts when we use R
- Report

The goal of this course

- I know what steps are involved in conducting (experimental) linguistic research!
- I know how PCIbex works! It doesn't mean I can set up any experiments, but I know where to start! I can at least also replicate studies such as a lexical decision task, and a self-paced reading task!
- I know how to plot a bar graph and a line graph using R!
- I know how to measure word similarity computationally, and calculate word/sentence surprisals, using libraries in Python!
- I know how to write up a research proposal incorporating all the things I learned!

The **goal** of this course

- This is our goal! So let’s give it a try!

Reminder

- This course targets students will different background.
- You come from Linguistics, Comp Sci, Cog Sci, and Data Science.
- You may find some assignments or readings easier or more challenging than others.
- Please help each other to fill in the missing gap! You're welcome to (and highly encouraged to) form a group of 2 for your final project.

Syllabus

- Syllabus
- R (download); Python (online – Google Colab); PCIbex (online)
- Sign up for Student demonstration.

Course website

1. https://tinyurl.com/exp-comp-methods-ling-research
   - Assignment instructions, and link to online tutorials will be posted on this website.

2. Canvas
   - Assignment submission, access to reading materials, discussion with your classmates

Next time

- No class this Tuesday – but please do your readings. These two readings will give you a good idea of this course.

- First assignment due next Thursday. These will be quite easy for those who have experience in R and Python. For those who are new to R and Python, please start working on it.