Computational Linguistics and Cognitive Science (Special Topics in Data Science)

NYU DS-GA 3001 003, Spring 2023
Tal Linzen (Linguistics and Data Science)
Wednesday 2:00-3:40 pm
60 5th Avenue, Room 204

Instructor	**Email**	**Office hour**
Tal Linzen	linzen@nyu.edu	Wednesdays, 11:30 am-12:30 pm 60 5th Avenue, Room 704 Sign up
Guy Davidson	gd1279@nyu.edu	Mondays, 2:15 PM - 3:15 PM 60 5th Avenue, Room 611

Course content

Overview

Humans are able to learn language rapidly and understand it robustly in the face of noise and variability. Recently, artificial intelligence systems have shown remarkable progress that appears to bring them closer to human-level language understanding. This class will take up two interrelated questions: First, how can we use computational simulations to understand humans' linguistic abilities and their neural bases? And second, how can we use cognitive science as a source of ideas and methods for further advancing artificial intelligence? The class assumes familiarity with contemporary natural language processing technologies (e.g. DS-GA 1011 or DS-GA 1012). Existing experience with computational cognitive modeling (e.g. DS-GA 1016) is recommended.

Schedule

Order of weeks is somewhat tentative and depends on the availability of guests.

Week 1 (1/25): Prediction and syntax

- Marten van Schijndel & Tal Linzen (2021). Single-stage prediction models do not explain the magnitude of syntactic disambiguation difficulty. *Cognitive Science*.
- Suhas Arehalli, Brian Dillon & Tal Linzen (2022). Syntactic surprisal from neural models predicts, but underestimates, human processing difficulty from syntactic ambiguities. *CoNLL*.

Week 2 (2/1): Syntactic dependencies

- Yair Lakretz et al. (2021). Mechanisms for handling nested dependencies in neural-network language models and humans. *Cognition*.
- Andrew Lampinen (2022). Can language models handle recursively nested grammatical structures? A case study on comparing models and humans. *arXiv*.

Week 3 (2/8): Syntactic transformations (Guest: Aaron Mueller, NYU)

- Ambridge, Rowland & Pine (2008). Is Structure Dependence an Innate Constraint? New Experimental Evidence From Children's Complex-Question Production. *Cognitive Science*.
- Aaron Mueller, Robert Frank, Tal Linzen, Luheng Wang & Sebastian Schuster (2022). Coloring the blank slate: Pre-training imparts a hierarchical inductive bias to sequence-to-sequence models. *Findings of ACL*.

Week 4 (2/15): Multimodal language learning
Deadline to identify a project topic.

- Chaz Firestone (2021). Performance vs. competence in human–machine comparisons. PNAS.
- Radford et al. (2021). Learning Transferable Visual Models From Natural Language Supervision. *ICML*. (Longer arXiv redux version.)
- Jack Merullo et al. (2022). Linearly Mapping from Image to Text Space. *arXiv*.

Week 5 (2/22): Grounded and distributional learning (Guest: Will Merrill, NYU)

- Arthur M. Glenberg & David A. Robertson (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language*.
- William Merrill, Alex Warstadt & Tal Linzen (2022). Entailment semantics can be extracted from an ideal language model. *CoNLL*.

Week 6 (3/1): Multimodal language learning (Guest: Wentao Wang, NYU
Proposal due.

- Wentao Wang, Wai Keen Vong, Najoung Kim & Brenden M. Lake (2022). Finding Structure in One Child's Linguistic Experience. *PsyArXiv*.
- Mitja Nikolaus, Afra Alishahi & Grzegorz Chrupała (2022). Learning English with Peppa Pig. *TACL*.

Week 7 (3/8)
Proposal presentations

(No class on 3/15/23: spring break)

Week 8 (3/22): Emergent compositionality

- Charles Lovering & Ellie Pavlick (2022). Unit Testing for Concepts in Neural Networks. *TACL*.
- Martha Lewis, Qinan Yu, Jack Merullo, Ellie Pavlick (2022). Does CLIP Bind Concepts? Probing Compositionality in Large Image Models. *arXiv*.

Week 9 (3/29): Emergent compositionality

- R. Thomas McCoy, Tal Linzen, Ewan Dunbar & Paul Smolensky (2019). RNNs implicitly implement tensor product representations. *ICLR*.
- Shikhar Murty, Pratyusha Sharma, Jacob Andreas, Christopher D. Manning (2022). Characterizing Intrinsic Compositionality in Transformers with Tree Projections. *arXiv*.

Week 10 (4/5): Composing word vectors (Guest: Allyson Ettinger, University of Chicago)
Progress report due

- Walter Kintsch (2001). Predication. *Cognitive Science*.

- Lang Yu & Allyson Ettinger (2020). Assessing Phrasal Representation and Composition in Transformers. *EMNLP*.
- Lalchand Pandia & Allyson Ettinger (2021). Sorting through the noise: Testing robustness of information processing in pre-trained language models. *EMNLP*.

Week 11 (4/12): Reasoning (Guest: Ishita Dasgupta, DeepMind)

- Johnson-Laird (1999). Deductive reasoning. *Cognition*.
- Ishita Dasgupta, Andrew K. Lampinen, Stephanie C. Y. Chan, Antonia Creswell, Dharshan Kumaran, James L. McClelland & Felix Hill (2022). Language models show human-like content effects on reasoning. *arXiv*.

Week 12 (4/19): Reasoning (Guest: Abulhair Saparov)

- Saparov & He (2023). Language Models Are Greedy Reasoners: A Systematic Formal Analysis of Chain-of-Thought. ICLR.
- Binz & Schultz (2023). Using cognitive psychology to understand GPT-3. PNAS.

Week 13 (4/26): Reasoning
Tal is traveling, but you still come in!

- Webb, Holyoak & Lu (2022). Emergent Analogical Reasoning in Large Language Models. *arXiv*.
- Khemlani & Johnson-Laird (2022). Reasoning About Properties: A Computational Theory. *Psychological Review*.

Week 14 (5/3)
Final project presentations

Course policies

Instruction format

This class will be taught in person. As this is a discussion-based course, regular attendance and participation are expected (see below). The class will not be recorded and it will not be possible to join the class via Zoom.

Lateness policy

All project-related assignments are due before the class starts (i.e. by 2 pm on Wednesday). The first late day (between one minute and 24 hours late) will result in a deduction of 5% of the grade for the assignment, and every additional late day will result in a deduction of 20% the grade. Students with relevant accommodations should discuss them with the instructor ahead of time.

Office hours

For Tal's office hours, please sign up for an appointment here earlier than Tuesday at 7 pm. Office hours are in person at 60 5th Avenue, room 704. Let me know if you have a conflict and need to meet at a different time.

Assessment

- Written responses to the readings (20%)
- Paper presentation (5%)
- Peer feedback on project proposals (5%)
- Attendance and participation (10%)
- Final project (55%)
    - Abstract (5% of final grade)
    - Proposal (15%)
    - Proposal presentation (5%)
    - Progress report (5%)
    - Final report (20%)
    - Final presentation (5%)
- Peer review for final project (5%)

Attendance and absences

This course depends on in-class discussion and participation, and students who miss a significant amount of class will have difficulty meeting the goals of the course. We therefore expect that students will make it their intention to attend every single class session consistently and on time.

You can miss one class without justification. Each unexcused missed class beyond the first one will lower your overall grade by three points, and if you are absent without excuse for three or more of the class meetings, your grade will be lowered a full letter grade (e.g., from a B+ to a C+). Religious observance and documented illness or family emergency are grounds for absences to be excused.

Laptop policy

Laptops are allowed in the classroom, but only for purposes relevant to the class. If you physically attend a class but spend the duration of the class doing your email, coding your homework for a different class, shopping for sweaters, etc, that will count as an absence for grading purposes.

Auditing

Auditors will be welcome to attend the lectures and will be provided access to the course materials, but will not be allowed to submit projects.

Written responses

Every week, you'll choose one of the papers assigned in that week, and submit the following four paragraphs about it: (1) a summary of the paper; (2) something you thought was a strength of the paper; (3) a weakness of the paper or a question you had about it; (4) an idea for a proposed follow-up experiment. The paragraphs don't need to be long, but they need to be concrete and specific. The responses should be written individually (not in teams), and are due by 7 pm the day before the class meets.

Paper presentations

Each student will present a paper in class. Make a 10-minute slide deck summarizing the motivation for the work, the methods and results of the main experiments of the paper, and the paper's takeaways. The goal of the presentation is to serve as a starting point for discussion in class: you can assume that the other students have read the paper, and you don't need to include all of the details and results in your presentations.

Course project

The project is expected to address a novel research question: it should not be a replication of existing work. The final report is expected to be at a level that can be submitted to an ACL workshop. Both solo and group projects are allowed, but the amount of work is expected to be proportional to the size of the group. You are expected to make some progress on the project most of the weeks of the semester, and are encouraged to meet with us throughout the semester to discuss it.

Topic and abstract: Please discuss the topic with us in our office hours as soon as possible. If you don't have an idea, sign up for a slot anyway, and we'll brainstorm together. The deadline to identify a topic is 2/15; by that date you should submit a 7-10 sentence abstract of the project you have in mind. You will receive feedback from the instructors within two days of the submissions deadline.

Proposal: The proposal should be 4 pages in ACL format. References are not included in the page count. The proposal has two main goals: first, review the relevant literature, with the goal of demonstrating that the question is novel; and second, describe the computational and/or human experiments you propose to address this question. The proposal is due on 3/1. You will receive feedback from the instructors within a week. You should start working on the project as soon as you get our feedback.

Proposal presentation: Each team will prepare a 10-minute oral presentation of its project proposal.

Feedback on proposal: In the two days following the presentation, all students are expected to post at least one suggestion or comment on each of the proposals to Brightspace. The comments should be as constructive as possible.

Progress report: On 4/5, about a month after the proposal is due, you will submit a report on the progress you've made on the project so far. This report can reuse text from the proposal, but must also document new empirical progress (e.g., dataset collection, experimental materials or results, etc.).

Final report: The report should be 8 pages in ACL format. References are not included in the page count. For group projects, the report must include a statement of the contribution of each team member. Deviations from the proposal are allowed, but only in case a research direction turned out to be unproductive. See below for writing tips.

Final presentation: Each team will prepare a 10-minute oral presentation of the final report and will present it in the last week.

Final project peer review: Each team will receive one other team's report, and will be expected to send to that team (and the instructors) a review of the report within a week. The review should be two pages in ACL format, and should include a summary of the paper, a few of its strengths, and a few constructive suggestions for improvement.

Writing tips for the final report

You're allowed to reuse prose from the proposal and progress report, but please make sure the paper reads as a coherent whole.

The report should be self-contained: it should describe the models and datasets you're using and hypotheses you're testing, without assuming that the reader already knows them, from the proposal or from anywhere else. The introduction should be largely understandable by a smart undergraduate student outside the field.

Clear writing is very important. I'm not going to take off points for grammar errors, but please read over your report carefully to make sure your sentences are concise, simple, and understandable by someone who is not closely familiar with your project. If your writing is difficult to follow, this may affect your grade even if the technical work you did was sound.

A typical report will consist of the following content:

1. An introduction motivating the project and describing prior work. Only cite a paper if it's relevant, for example, if you want to explain how your method or research question differs from existing work.
2. Methods (or Experimental Setup): your models, your datasets and how you ran your experiments.
3. Results: Include figures and/or tables. Figures are usually better than tables. All figures and tables should have a caption, which should be as self-contained as possible (avoid “see text for details”). Text in the figures should be readable without zooming in.
4. Discussion: Was your hypothesis confirmed? How do your findings relate to findings from prior work? What are the limitations of your experiment, and how can they be addressed in future work?
5. Conclusion: One or two paragraphs summarizing the main takeaway.
6. Author contribution statement (for team projects).

Relevant university policies

Academic Integrity

Work you submit should be your own. Please consult the CAS academic integrity policy for more information: https://cas.nyu.edu/content/nyu-as/cas/academic-integrity.html – penalties for violations of academic integrity may include failure of the course, suspension from the University, or even expulsion.

Religious Observance

As a nonsectarian, inclusive institution, NYU policy permits members of any religious group to absent themselves from classes without penalty when required for compliance with their religious obligations. The policy and principles to be followed by students and faculty may be found here: The University Calendar Policy on Religious Holidays (http://www.nyu.edu/about/policies-guidelines-compliance/policies-and-guidelines/university-calendar-policy-on-religious-holidays.html)

Disability Disclosure Statement

Academic accommodations are available for students with disabilities. The Moses Center website is www.nyu.edu/csd. Please contact the Moses Center for Students with Disabilities (212-998-4980 or mosescsd@nyu.edu) for further information. Students who are requesting academic accommodations are advised to reach out to the Moses Center as early as possible in the semester for assistance.