Lsci 209: Seminar on Information Theory and Language
Tuesdays and Thursdays 2-3:20p, SSL 152
Instructor: Richard Futrell (rfutrell@uci.edu)
Office hours: M 10am in SSPB 2215 or http://uci.zoom.us/my/futrell or by appointment

Course Content. Information theory is the mathematical theory of communication, relating the theory of probability to the theory of codes. It has applications to Language Science in at least four areas:

(1) In explaining properties of languages when we view human languages as information-theoretic codes operating under certain constraints.
(2) In characterizing linguistic structure statistically and probabilistically.
(3) In characterizing human online language processing (comprehension and production) and linguistic behavior (pragmatics).
(4) In formulating, interpreting, and understanding machine learning methods for natural language technologies, especially those based on neural networks.

An additional area where information theory is promising, but which is not as well-developed in the literature, is in developing theories of language acquisition.

We will read mostly recent papers in these areas, primarily from the last three years, from conferences and journals on linguistics, cognitive science, neuroscience, natural language processing, and information theory. These papers are highly technical. I can provide technical details and background as needed. Our goal is to extract the core ideas that might be applicable to the problems we are all interested in. Readings will also include some things that are not explicitly about either information theory or language (for example, several cognitive neuroscience readings toward the end of the class). The goal for these readings is to try to forge a connection which might not be there in the paper itself, but where I think one is possible.

Course Structure. We will meet twice a week and discuss a paper or group of papers at each meeting. Students will prepare summary presentations on papers and lead discussions by preparing discussion questions.

Final Project. You will complete a final project for the class. You will turn in a writeup and give a presentation to the class so we can all discuss. Presentations will be in the final week of class and the writeup will be due at the end of the finals period. For more information on the final project, see the end of this document.

Grading. Your final grade is 50% final project, 25% presentation of 2 papers, and 25% participation in discussions in class.

Schedule (subject to modification)

Date	Content	Who presents?
T 1/4	–	
R 1/6	*Planning, Organization, and Big Questions*	
T 1/11	Edward Gibson, Richard Futrell, Steven T. Piantadosi, Isabelle Dautriche, Kyle Mahowald, Leon Bergen, and Roger P. Levy (2019). How efficiency shapes human language. *Trends in Cognitive Sciences.*	Shiva
R 1/13	Lexical Efficiency Tiago Pimentel, Rowan Hall Maudslay, Damián Blasi, and Ryan Cotterell (2020). Speakers fill lexical semantic gaps with context. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 4004-4015. Tiago Pimentel, Irene Nikkarinen, Kyle Mahowald, Ryan Cotterell, and Damián Blasi (2021). How (Non-)Optimal is the Lexicon? In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4426–4438.	*Context:* Zeinab *(Non-)Optimal:* Kolby
T 1/18	Information Bottleneck for Lexical Semantics Noga Zaslavsky, Charles Kemp, Terry Regier, and Naftali Tishby (2018). Efficient compression in color naming and its evolution. *Proceedings of the National Academy of Sciences* 115(31): 7937-7942.	Catarina
R 1/20	Information Bottleneck for Grammatical Marking Francis Mollica, Geoff Bacon, Noga Zaslavsky, Yang Xu, Terry Regier, and Charles Kemp (2021). The forms and meanings of grammatical markers support efficient communication. *Proceedings of the National Academy of Sciences* 118(49): e2025993118. Noga Zaslavsky*, Mora Maldonado*, and Jennifer Culbertson (2021). Let's talk (efficiently) about us: Person systems achieve near-optimal compression. In	*Grammatical markers:* Weijie *Person systems:* Charlie

	Proceedings of the 43rd Annual Meeting of the Cognitive Science Society.	
T 1/25	Surprisal Theory Nathaniel J. Smith and Roger P. Levy (2013). The effect of word predictability on reading time is logarithmic. *Cognition* 128(3): 302–319. Danny Merkx and Stefan L. Frank (2021). Human Sentence Processing: Recurrence or Attention? In *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics*, pages 12–22.	*Smith:* Richard *Merkx:* Jiaxuan
R 1/27	Uniform Information Density (UID) T. Florian Jaeger (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology* 61(1): 23–62.	Niels
T 2/1	Further UID Clara Meister, Tiago Pimentel, Patrick Haller, Lena Jäger, Ryan Cotterell, and Roger P. Levy (2021). Revisiting the Uniform Information Density Hypothesis. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 963–980.	Yongjia
R 2/3	Information Rate Christophe Coupé, Yoon Mi Oh, Dan Dediu, and François Pellegrino (2019). Different languages, similar encoding efficiency: Comparable information rates across the human communicative niche. *Science Advances* 5(9): eaaw2594. Tiago Pimentel, Clara Meister, Elizabeth Salesky, Simone Teufel, Damián Blasi, and Ryan Cotterell (2021). A surprisal–duration trade-off across and within the world's languages. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 949–962.	*Coupé:* Shivanshu *Pimentel:* Noa

T 2/8	Words and Idioms John Mansfield (2021). The word as a unit of internal predictability. *Linguistics.* Michaela Socolof, Jackie Chi Kit Cheung, Michael Wagner, and Timothy J. O'Donnell (2021). Characterizing Idioms: Conventionality and Contingency. *arXiv* 2104.08664.	*Mansfield:* Manikanta *Socolof:* Yongjia
R 2/10	Syntax Richard Futrell, Peng Qian, Edward Gibson, Evelina Fedorenko, and Idan Blank (2019). Syntactic dependencies correspond to word pairs with high mutual information. In *Proceedings of the Fifth International Conference on Dependency Linguistics*, pages 3-13. Jacob Louis Hoover, Wenyu Du, Alessandro Sordoni, and Timothy J. O'Donnell (2021). Linguistic dependencies and statistical dependence. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 2941–2963.	*Futrell:* Jiaxuan *Hoover:* Catarina
T 2/15	Information-Theoretic Probing of Neural Networks Elena Voita and Ivan Titov (2020). Information-Theoretic Probing with Minimum Description Length. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 183–196. Tiago Pimentel and Ryan Cotterell (2021). A Bayesian Framework for Information-Theoretic Probing. *EMNLP 2021*.	*Voita:* Shivanshu *Pimentel:* Kolby
R 2/17	Language Acquisition Francis Mollica and Steven T. Piantadosi. Humans store about 1.5 megabytes of information during language acquisition. *Royal Society Open Science* 6(3): 181393.	Weijie
T 2/22	Cognitive Neuroscience	*Lai:* Shiva

	Lucy Lai and Samuel J. Gershman. Policy compression: An information bottleneck in action selection. *Psychology of Learning and Motivation* 74: 195–232. Alexandre Zénon, Oleg Solopchuk, and Giovanni Pezzulo (2019). An information-theoretic perspective on the costs of cognition. *Neuropsychologia* 123: 5–18.	*Zénon:* Niels
R 2/24	Free Energy Principle Karl Friston (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience* 11(2): 127–138. Samuel J. Gershman (2019). What does the free energy principle tell us about the brain? *Neurons, Behavior, Data Analysis, and Theory*.	*Friston:* Charlie *Gershman:* Zeinab
T 3/1	Reinforcement Learning / Social Dynamics Natasha Jaques, Angeliki Lazaridou, Edward Hughes, Caglar Gulcehre, Pedro A. Ortega, DJ Strouse, Joel Z. Leibo, and Nando de Freitas (2019). Social Influence as Intrinsic Motivation for Multi-Agent Deep Reinforcement Learning. In *Proceedings of the 36th International Conference on Machine Learning*. Rahma Chaabouni, Eugene Kharitonov, Diane Bouchacourt, Emmanuel Dupoux, and Marco Baroni (2020). Compositionality and Generalization In Emergent Languages. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 4427-4442.	*Jaques:* Manikanta *Chaabouni:* Noa
R 3/3	*Student Presentations*	Weijie Charlie Niels Noa
T 3/8	*Student Presentations*	Catarina Shiva Shivanshu Jiaxuan
R 3/10	*Student Presentations*	Manikanta Kolby

		Yongjia Zeinab
F 3/18 5pm	*Final Project Writeup Due on Canvas*	

Final Project

You will complete a final project for the class. The final project is intended to enable you to bring together some of the material we discussed this term. Ideally it should be related to your own research. It is ok and encouraged for you to use your own ongoing research as material for the project.

You will need to select an area of interest, then identify a particular problem you might investigate in this area. One use of this assignment is to allow you to think about larger projects that you might later pursue, or to perhaps formalize an idea which you may go on and complete later—so please interpret these requirements liberally, and feel free to contact me if you are not sure of a project's suitability.

You should prepare a presentation to give in the last week of class, and then turn in a writeup of 12ish pages double-spaced where you:

1. Set up the problem
    1. Identify an area that interests you. The topic should be amenable to investigation using one of the broad range of methods (e.g. experiments, modeling) we looked at in class.
    2. Write a brief survey outlining the background of the area you have chosen.
    3. Identify a question within the area you have chosen, explain what the question is, then construct and describe a hypothesis through which you aim to address this question.
2. Design an investigation to test your idea.
    1. Describe and explain the design of a project (e.g. an experiment, whether computational or with human subjects) and how it aims to address your question.
    2. Specify what predictions you have formed in relation to your design.
3. Analysis
    1. If you actually carry out your project, analyze the results. If your project is just a design, describe how you would analyze your results.
    2. What criteria would you use to establish success / failure of your project.
4. Discussion
    1. Assuming that your project found what you expected, discuss this outcome in terms of the effect it would have on the broader area in which your research is situated, and what future research questions it may suggest.
    2. What if the results of your project were not what you expected, what would that mean? (If your project is a design -- i.e., if you don't have any results -- you may find it a useful exercise to write two discussion sections.)

Pandemic information

When physically present in a classroom, other instructional space, or any other space owned or controlled by UCI, all students and all employees (faculty and staff) must comply with COVID-19 related UCI executive directives and guidance. This guidance takes into account federal, state, and local guidelines and is available at https://uci.edu/coronavirus/.

Academic Integrity

We will be adhering fully to the standards and practices set out in UCI's policy on academic integrity. Any attempts of academic misconduct or plagiarism will be met with consequences as per the university regulations.

Disability

Any student requesting academic accommodations based on a disability is required to apply with Disability Service Center at UCI. For more information, please visit http://disability.uci.edu/.