

# Syllabus

## Computational Psycholinguistics

Psych/CogSt 4280/6280 - Ling 4428/6628

Fall 2015

205 Uris Hall - Tuesdays, 1:25-4:00PM

<http://courses.cit.cornell.edu/psych428/>

**Instructor:** Dr. Morten H. Christiansen

**Office:** 228 Uris Hall

**Email:** [christiansen@cornell.edu](mailto:christiansen@cornell.edu)

**Office Hours:** Tuesdays 12Noon-1:00PM and by appointment

Computational methods have become increasingly important to the study of the psychology of language as a way of gaining insights into the psychological processes involved in language processing, acquisition, and evolution. They provide rigorous tools for testing and exploring specific hypotheses about the nature of language and its psychological underpinnings. As such, computational psycholinguistics has already had a far-reaching impact on language research. In this course, we survey the state of the art in computational psycholinguistics, ranging from corpus analyses of child-directed speech, dialogues and online interactions, to Bayesian, connectionist and phylogenetic models of language. An important focus of discussion will be the various methodological and theoretical issues relating to the use of computational tools to understand the psychology of language.

**Learning goals:** To provide a comprehensive introduction to questions, theories, and research in computational psycholinguistics. At the end of this seminar, students are expected to be able to think critically about research and theories related to computational approaches to psycholinguistic phenomena.

### Course Requirements

1. Complete **ALL** assigned readings.
2. Lead discussion on one or more occasions (the exact number will depend on number students registered for the class). Joint presentations are preferable and encouraged.
3. Discussion leaders should only briefly summarize the reading(s). The leader should assume that everyone has read the reading(s), and does not need to have it repeated in great detail. The discussion should focus on your elaborations of the reading(s). This involves clarifying the reading(s) (guided in part by questions submitted by other class members), critiquing the research, and *including other material/viewpoints from additional articles*.
4. Formulate at least one question for each assigned reading, for a total of three questions for the day's discussion leader. These are to be submitted by email to the presenter **before 10AM on Mondays before class**. A copy of the questions must also be emailed to the instructor ([christiansen@cornell.edu](mailto:christiansen@cornell.edu)).
5. Participate in the discussions. Grades will in part be based on discussion participation. Sitting silently through every class is **not** acceptable.
6. Write a final 15-20-page (4280/4428) or 25-30-page (6280/6628) double-spaced paper focusing on several of the topics discussed in class. This paper will require some research, and must include additional articles not discussed in class. A one-page synopsis outlining a proposed paper must be submitted to the instructor at the start of class November 10. The final paper is due December 11 before 4:30PM.

### Grading

Grades will be based on class presentations, discussion participation, email questions, and the final paper.

Class presentation	20%
Discussion participation	25%
Email questions	25%
Final paper	30%

### Academic Integrity

Each student in this course is expected to abide by the Cornell University Code of Academic Integrity. Any work submitted by a student in this course for academic credit will be the student's own work. For this course, collaboration is allowed in the following instances: Leading discussions. Failure to adhere to the Code of Academic Integrity will result in an F in the course.

## Course Outline

\* indicates extra readings for presenters and 6280/6628 students

<b>Week</b>	<b>Dates</b>	<b>Topics</b>	<b>Readings</b>
Week 1	8/25	Organizational meeting	
Week 2	9/1	Psycholinguistics and computational modeling	Bernstein Ratner, Berko Gleason & Narasimhan (1998) Chater & Christiansen (2008)
Week 3	9/8	Segmentation	Phillips & Pearl (2015) Baayen et al. (in press) *French & Cottrell (2014)
Week 4	9/15	<i>No class</i>	
Week 5	9/22	Word learning	McMurray et al. (2012) Hills et al. (2009) *Monaghan et al. (2011)
Week 6	9/29	Semantic representations	Blouw et al. (2015) Xu et al. (2015) *Youn et al. (2015)
Week 7	10/6	Zipf's law	Piantadosi (2014) Ferrer-i-Cancho et al. (2013) *Pine et al. (2013)
	10/13	<b>Fall Break</b>	
Week 8	10/20	Chunking	Jones (2015) McCauley & Christiansen (2014a) *Kol et al. (2014)
Week 9	10/27	Politeness and the linguistic change in online communities  <i>Guest lecture: Dr. Cristian Danescu-Niculescu-Mizil</i>	Danescu-Niculescu-Mizil et al. (2013a) Danescu-Niculescu-Mizil et al. (2013b)
Week 10	11/3	Acquisition of morphology	Xanthos et al. (2012) Mirković et al. (2011) *Aguado-Orea & Pine (2015)
Week 11	11/10	Sentence processing  <b>1-page synopsis due</b>	Johns & Jones (2015) Frank & Bod (2011) *McCauley & Christiansen (2014b)

Week	Dates	Topics	Readings
Week 12	11/17	Production	Dell et al. (2013) Montag & MacDonald (2015) *Arnon & Cohen-Priva (2013)
Week 13	11/24	Dialogue	Fusaroli & Tylén (2015) Abney et al. (2014) *Healey et al. (2014)
Week 14	12/1	Evolution	Kirby et al. (2015) Baronchelli et al. (2012) *Grouchy et al. (2015)
	12/11	<i>Final paper due</i> (before 4:30PM)	

**Note:** Changes may be made to the readings and their order but this will be announced in class and on the course web site.

### Readings

- Abney, D. H., Paxton, A., Dale, R., & Kello, C. T. (2014). Complexity matching in dyadic conversation. *Journal of Experimental Psychology: General*, *143*, 2304-2315.
- \*Aguado-Orea, J. & Pine, J.M. (2015). Comparing different models of the development of verb inflection in early child Spanish. *PLoS ONE* *10*(3): e0119613.
- \*Arnon, I., & Cohen Priva, U. (2013). More than words: The effect of multi-word frequency and constituency on phonetic duration. *Language and Speech*, *56*, 349-371.
- Baayen, R. H., Shaoul, C., Willits, J. and Ramscar, M. (in press). Comprehension without segmentation: A proof of concept with naive discrimination learning. *Language, Cognition, and Neuroscience*.
- Baronchelli, A., Chater, N., Pastor-Satorras, R. & Christiansen, M.H. (2012). The biological origin of linguistic diversity. *PLoS ONE*, *7*(10): e48029.
- Bernstein Ratner, N., Berko Gleason, J. & Narasimhan, B. (1998). Introduction to psycholinguistics: What do language users know? In J. Berko Gleason & N. Bernstein Ratner (Eds.), *Psycholinguistics* (2<sup>nd</sup> ed., pp. 1-40). Orlando, FL: Harcourt Brace.
- Blouw, P., Solodkin, E., Thagard, P., & Eliasmith, C. (2015). Concepts as semantic pointers: A framework and computational model. *Cognitive Science*.
- Chater, N. & Christiansen, M.H. (2008). Computational models in psycholinguistics. In R. Sun (Ed.), *Cambridge Handbook of Computational Cognitive Modeling* (pp. 477-504). New York: Cambridge University Press.
- Danescu-Niculescu-Mizil, C., Sudhof, M., Jurafsky, D., Leskovec, J., & Potts, C. (2013a). A computational approach to politeness with application to social factors. *arXiv preprint arXiv:1306.6078*.
- Danescu-Niculescu-Mizil, C., West, R., Jurafsky, D., Leskovec, J., & Potts, C. (2013b). No country for old members: User lifecycle and linguistic change in online communities. In *Proceedings of the 22nd international conference on World Wide Web* (pp. 307-318). International World Wide Web Conferences Steering Committee.
- Dell, G. S., Schwartz, M. F., Nozari, N., Faseyitan, O., & Coslett, H. B. (2013). Voxel-based lesion-parameter mapping: Identifying the neural correlates of a computational model of word production. *Cognition*, *128*, 380-396.
- Ferrer-i-Cancho, R., Hernández-Fernández, A., Lusseau, D., Agoramoorthy, G., Hsu, M. J. & Semple, S. (2013). Compression as a universal principle of animal behavior. *Cognitive Science*, *37*, 1565-1578.
- Frank, S.L. & Bod, R. (2011). Insensitivity of the human sentence-processing system to hierarchical structure. *Psychological Science*, *22*, 829-834.
- \*French, R. & Cottrell, G.W. (2014) TRACX 2.0: A memory-based, biologically-plausible model of sequence segmentation and chunk extraction. In *Proceedings of the 36th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.
- Fusaroli, R., & Tylén, K. (2015). Investigating conversational dynamics: Interactive alignment, Interpersonal synergy, and collective task performance. *Cognitive science*.

- \*Grouchy, P.E., D'Eleuterio, G.M.T, Lipson, H. & Christiansen, M.H. (2015). *The evolutionary origin of symbolic communication*. Submitted manuscript.
- \*Healey, P.G.T., Purver, M. & Howes, C. (2014). Divergence in dialogue. *PLoS ONE* 9(6): e98598.
- Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009). Longitudinal analysis of early semantic networks: Preferential attachment or preferential acquisition?. *Psychological Science*, 20, 729-739.
- Johns, B. T., & Jones, M. N. (2015). Generating structure from experience: A retrieval-based model of language processing. *Canadian Journal of Experimental Psychology*.
- Jones, G. (2015). *The influence of children's exposure to language from two to six years: The case of nonword repetition*. Submitted manuscript.
- Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, 141, 87-102.
- \*Kol, S., Nir, B., & Wintner, S. (2014). Computational evaluation of the Traceback Method. *Journal of child language*, 41, 176-199.
- McCauley, S.M. & Christiansen, M.H. (2014a). Acquiring formulaic language: A computational model. *Mental Lexicon*, 9, 419-436.
- \*McCauley, S.M. & Christiansen, M.H. (2014b). Prospects for usage-based computational models of grammatical development: Argument structure and semantic roles. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5, 489-499.
- McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review*, 119, 831-877.
- Mirković, J., Seidenberg, M. S., & Joanisse, M. F. (2011). Rules versus statistics: Insights from a highly inflected language. *Cognitive Science*, 35, 638-681.
- \*Monaghan, P., Christiansen, M.H. & Fitneva, S.A. (2011). The arbitrariness of the sign: Learning advantages from the structure of the vocabulary. *Journal of Experimental Psychology: General*, 140, 325-347.
- Montag, J. L., & MacDonald, M. C. (2015). Text exposure predicts spoken production of complex sentences in 8- and 12-year-old children and adults. *Journal of Experimental Psychology: General*, 144, 447-468.
- Phillips, L. & Pearl, L. (2015). Utility-based evaluation metrics for models of language acquisition: A look at speech segmentation. *Workshop on Cognitive Modeling and Computational Linguistics 2015*, NAACL.
- Piantadosi, S. T. (2014). Zipf's word frequency law in natural language: A critical review and future directions. *Psychonomic bulletin & review*, 21, 1112-1130.
- \*Pine, J. M., Freudenthal, D., Krajewski, G., & Gobet, F. (2013). Do young children have adult-like syntactic categories? Zipf's law and the case of the determiner. *Cognition*, 127, 345-360.
- Xanthos, A., Lahaa, S., Gillis, S., Stefany, U., Aksu-Koc, A., Christofidou, A., Gagarina, N., Hrzica, G., Nihan Ketrez, F., Kilani-Schoch, M., Korecky-Kroll, K., Kovacevic, M., Laalo, K., Palmovic, M., Pfeiler, B., Voeikova, M. D. & Dressler, W. U. (2012). On the role of morphological richness in the early development of noun and verb inflection. *First Language*, 31, 461-479.
- Xu, Y., Regier, T., & Malt, B.C. (2015). Semantic chaining and efficient communication: The case of container names. In *Proceedings of the 37<sup>th</sup> Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.
- \*Youn, H., Sutton, L., Smith, E., Moore, C., Wilkins, J. F., Maddieson, I., ... & Bhattacharya, T. (2015). On the universal structure of human lexical semantics. *arXiv preprint arXiv:1504.07843*.