Syllabus: Variation in phonology

Time/Place   TR 4:00-5:50, Campbell Hall 2101K

Course web site: http://www.linguistics.ucla.edu/people/zuraw/251_2013

	Bruce Hayes	**Kie Zuraw [ˈkʰaj ˈzɻ̩ˌɔ]**
Mailbox	In Campbell 3125	In Campbell 3125
Phone	310-825-9507	310-825-0634
Student hrs.	W,F 2-3 (tentative)	TBA
	Campbell 2101G	Campbell 3122A
Email	bhayes@humnet.ucla.edu	kie@ucla.edu
Web	linguistics.ucla.edu/people/hayes	linguistics.ucla.edu/people/zuraw

Description and goals

Variation pervades phonological behavior. In recent years phonologists have begun rising to the challenge of developing theories for the treatment of variation. We (Hayes and Zuraw) are interested in the topic as a theoretical problem — but also as a way to advance research in phonology on all fronts, by helping us to construct grammars that make explicit predictions about the data, including its pattern of variation.

The goals of this course are twofold.

- First, we want to address the theoretical question of how variable patterns are learned and used, and how data can be used to make arguments bearing on these issues.

- Second, at a more practical level, we want to equip you with theoretical and practical tools (particularly, software) that can help you deal with variation in your own data. The goal is for you to be able to use these tools to model your data, and to use your data to argue for or against a model.

Course requirements

- For four units:
    - ➢ Do the readings, some of which will involve writing a less-than-page outline/summary
    - ➢ Do the software exercises: mostly little things where you learn to use software packages to do analysis
    - ➢ Write a term paper that bears on the course content. You need to have the topic approved by talking to one or both of us.

- For two units:
    - ➢ Do the readings (honor system)
    - ➢ Do two of the exercises (you choose)

Some planned exercises

- apply GLA to a dataset
- apply Noisy Harmonic Grammar to a dataset
- apply Maxent to a dataset
- apply regression (logistic, linear) to a dataset
- a homework on model comparison

We're also interested in skills for data-crunching (study of electronic corpus data) and may have a lab on this.

Course outline

- What sorts of variation exist? *(Coetzee & Pater 2011 for overview)*
  - Free variation vs. lexical variation (and mixed variation)
  - Type (lexical) vs. token (in production)
  - Overview of token variation: the classical sociolinguistic tradition (Cedergren & Sankoff 1974)
  - Overview of type variation: the Law of Frequency Matching, studies supporting it; the nuances and deviations
- Theories
  - Probability distributions over classic OT grammars
    - Freely-ranked strata: *Anttila 1997*
    - Stochastic OT: *Boersma & Hayes 2001*
  - Reconciling constraint preferences through weights rather than strict ranking
    - Noisy Harmonic Grammar *(Coetzee 2009)*
    - Maxent *(Goldwater & Johnson 2003)*
  - MaxEnt's relation to logistic regression (introduces important concepts like loss function/measure of fit, significance)
  - How to decide on these models? Hayes/Zuraw work in progress.
- The problem of model selection: how to know which model is (closer to) correct?
  - Levels of model selection:
            grammatical architecture
            constraint set
            parameter values, e.g. weights or even the priors
  - Methods of model selection:
            statistical approaches
            machine learning approaches (cross-validation)
- Biases in UG and priors in modeling
  - The types of bias: naturalness, simplicity, limitations on constraint weight
  - underfitting and overfitting
  - priors and smoothing in statistics and machine learning
  - linguistic results with priors: *Wilson 2006, Ryan 2010, Martin 2011, White 2012, Hayes/Zuraw et al. 2011*
- Letting the grammar output nothing

- ➢ This actually happens (??*fishish*) — when we do it stochastically, we get a measure of productivity. *Steriade's -able, Raffelsiefen*
- ➢ using the models with the Null Parse
- ➢ Hay and Baayen's work
- ➢ Lexical selection and filters in general *(Martin 2007)*
- ➢ filters and goodness scores—not all winning candidates are equally good

- Gradient intuitions
  - ➢ Extending the scope of the variation models to intuition
  - ➢ Essential ingredient: the mapping of probability to ratings data

- Serialist models and variation
  - ▪ Level ordering as the basis of variation *(Guy 1991; Myers 1995; Kiparsky 1994)*
  - ▪ Harmonic serialism *(Kimper 2011)*
  - ▪ (McPherson & Hayes 2012)

- Exemplar models
  - ▪ *Sloos 2013*

- Multi-site variation *(Kaplan 2011's typology)*

Anttila, Arto. 1997. Deriving variation from grammar.. In Frans Hinskens, Roeland van Hout, & W. Leo Wetzels (eds.), *Variation, Change, and Phonological Theory*, 35–68. Amsterdam: John Benjamins.

Boersma, Paul & Bruce Hayes. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32. 45–86.

Cedergren, Henrietta J. & David Sankoff. 1974. Variable Rules: Performance as a Statistical Reflection of Competence. *Language* 50(2). 333–355. doi:10.2307/412441 (18 April, 2012).

Coetzee, Andries W. 2009. An integrated grammatical/non-grammatical model of phonological variation.. In Young-Se Kang, Jong-Yurl Yoon, Hyunkung Yoo, Sze-Wing Tang, Yong-Soon Kang, Youngjun Jang, Chul Kim, Kyoung-Ae Kim, & Hye-Kyung Kang (eds.), *Current Issues in Linguistic Interfaces*, vol. 2, 267–294. Seoul: Hankookmunhwasa.

Coetzee, Andries W & Joe Pater. 2011. The place of variation in phonological theory.. In John A Goldsmith, Jason Riggle, & Alan C. L. Yu (eds.), *The Handbook of Phonological Theory*, 401–434. John Wiley & Sons.

Goldwater, Sharon & Mark Johnson. 2003. Learning OT Constraint Rankings Using a Maximum Entropy Model.. In Jennifer Spenader, Anders Eriksson, & Östen Dahl (eds.), *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*, 111–120. Stockholm: Stockholm University.

Guy, Gregory R. 1991. Explanation in Variable Phonology: An Exponential Model of Morphological Constraints. *Language Variation and Change* 3(01). 1–22. doi:10.1017/S0954394500000429.

Kaplan, Aaron F. 2011. Variation Through Markedness Suppression. *Phonology* 28(03). 331–370. doi:10.1017/S0952675711000200.

Kimper, Wendell A. 2011. Locality and globality in phonological variation. *Natural Language & Linguistic Theory*.

Martin, Andrew. 2007. The evolving lexicon.. University of California, Los Angeles ph.d. dissertation.

Martin, Andrew. 2011. Grammars leak: modeling how phonotactic generalizations interact within the grammar. *Language* 87(4). 751–770.

McPherson, Laura & Bruce Hayes. 2012. Relating application frequency to morphological structure: the case of Tommo So vowel harmony.. Paper presented at the Manchester Phonology Meeting.

Myers, James. 1995. The categorical and gradient phonology of variable t-deletion in English.. Paper presented at the International Workshop on Language Variation and Linguistic Theory, University of Nijmegen, ms.

Ryan, Kevin M. 2010. Variable affix order: grammar and learning. *Language* 86(4). 758–791. (13 April, 2012).

Sloos, Marjoleine. 2013. *Phonological Grammar and Frequency: an Integrated Approach: Evidence from German, Indonesian and Japanese.*

White, James. 2012. Evidence for a learning bias against "saltatory" phonological alternations in artificial language learning.. Paper presented at the Linguistic Society of America Annual Meeting, Portland, OR.

Wilson, Colin. 2006. Learning Phonology with Substantive Bias: An Experimental and Computational Study of Velar Palatalization. *Cognitive Science* 30(5). 945–982.