LSA Institute 2015
Course 308
July 20, 2015

Day 1: Sound change and computational models

1 Administrative

- Go over syllabus
- Technical matters
- Office hours
- For Thursday: Read
  - Pierrehumbert (2001), Garrett (2014)
  - Optional: Wedel (2004), McElreath and Boyd (2007) Chapter 1

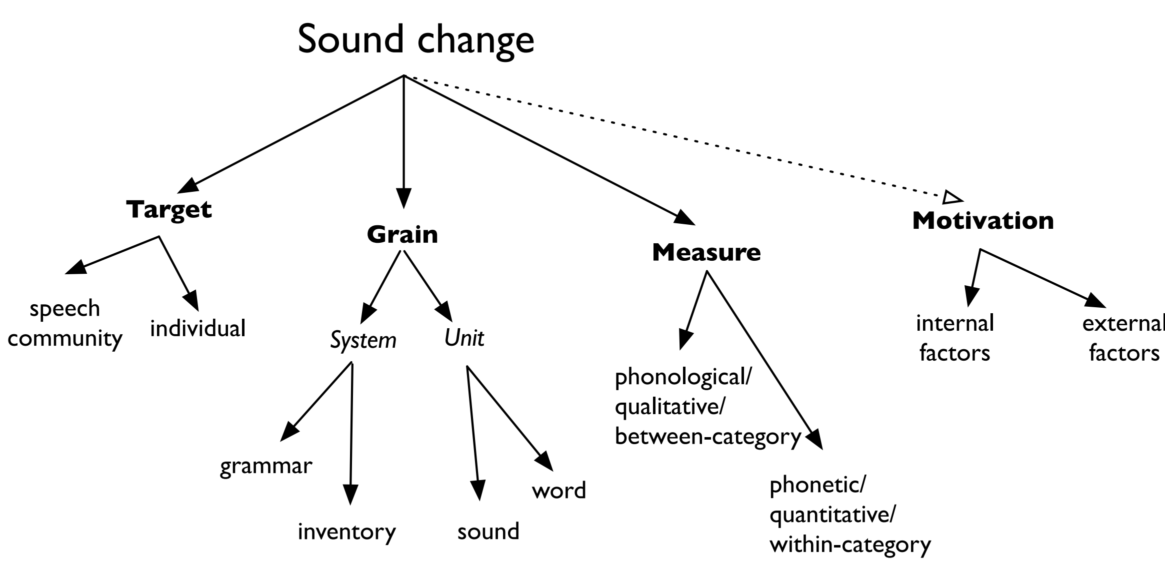
Today:

1. What is sound change?
2. Why model?
3. How do we model?

2 What is sound change?

What are we modeling?

- **'Sound change' means different empirical phenomena to different people/groups**[1]
  - Dimensions: Target, grain, measure (also: motivation)



- **Examples**

[1] "... there is no generally accepted definition of sound change" (Garrett, 2014)

1. Neogrammarian segmental change in historical linguistics (e.g. /b/→/v/, many languages; Campbell, 1998)
   - Sounds, community, qualitative
2. Lexical diffusion (Wang, 1969; Phillips, 1984)
   - Words, community, qualitative
3. Chain shifts in progress in English vowel systems (Hay et al., 2015; Labov, 1994, 2000)
   - Inventory, community, qualitative or quantitative
4. Vowel mergers in progress in English dialects (Hay et al., 2006)
   - Inventory, community, quantitative
5. Pronunciation change over the lifespan (years) (Harrington et al., 2000; Sankoff, 2005)
   - Sounds, individuals, quantitative/qualitative
6. Word-specific phonetic change (Pierrehumbert, 2001; Wedel, 2006)
   - Words/sounds, individuals, quantitative
7. etc...

- **Constituencies**
  1. (classic) Historical linguists
     - Long tradition; detailed taxonomy of sound changes (e.g. Hock, 1991)
     - Sound change *itself* rarely of interest; used as a tool to get at relatedness and change.
  2. Sociolinguists, sociophoneticians
     - Focus on (variation and) change in their own right: mergers, chain shifts, etc.
     - Differ in specific foci (e.g. 'classic' socio examines more qualitative variables, sociophonetics more quantitative variables)
  3. Phoneticians, phonologists
     - Both interested in *phonologization* (Beddor, Ohala, Bermudez-Otéro)
     - Grammar change/'optimization' (Kiparsky, Boersma, Flemming)
- **Common taxonomy of subproblems of language change** (Weinreich et al., 1968)
  - *Actuation**: Why do structural changes take place in a particular language at a given time, but not in other languages with the same feature, or in the same language at other times?
  - *Constraints**: What is the set of possible changes and possible conditions for change?
  - *Transition*: What is the route by which a linguistic change is proceeding to completion? What happens when these routes diverge?
  - *Embedding*: How is sound change embedded in linguistic/social structure? What other changes are associated with the given changes in a manner that cannot be attributed to chance?
  - *Evaluation*: What are the effects of the change on linguistic structure?

- **When/where does transmission take place?**
  1. L1 aquisition (mistransmission, imperfect learning)
     - 'Acquisitionism': change *only* occurs in L1 acquisition (widespread tacit assumption?)
     - Labov (1994, 2007): "generational change" is the norm for sound change internal to a speech community: "transmission" and "incrementation" in childhood/adolescence
  2. Change over the lifespan (e.g. Harrington et al., 2000; Sankoff, 2005)
     - The Change By Accommodation model (see Auer and Hinskens, 2005): Neogrammarians on
       (a) Imitation/accommodation/convergence
       (b) Change in individual
       (c) Change in community

3 Why model?

What does modeling contribute, as opposed to e.g. theoretical, laboratory approaches?

3.1 Advantages of modeling

- Adapted from McElreath and Boyd (2007):
  1. **Existence proofs.** *Could the model in principle account for the observed behaviour?*
     - There are many possible accounts of any pattern observed in a social or biological system
     - Modelling can help us determine what is a *possible* explanation
     - Often impossible to access relevant data we need/want
     - "Hazards of unaided reasoning"[2]
  2. **Communication.** *Explicit models are reproduceable and extendable.*
     - Allows us to explore the ramifications of our assumptions
     - Just because two models share assumptions $\neq$ they will produce the same results
  3. **Counterintuitive results.** *Iterating weak forces over time generally leads to surprises.*
     - Not always obvious how different factors will interact
     - Intuitions based on purely verbal descriptions can often be misleading
  4. **(Qualitative) predictions.** *Even simple models let us make qualitative predictions.*
     - "An increase in population size should increase the probability of merger"
     - "The more noise in the system, the smaller the predicted size of the stable inventory"
     - Models can also generate new hypotheses that can then be investigated empirically

[2] "With purely verbal arguments about evolutionary processes, it is too often the case that our conclusions do not follow from our assumptions. Unaided reasoning about the mass effects of many small forces operating over many generations has proven to be hazardous. Formalizing our arguments helps us understand which stories are possible explanations." (McElreath and Boyd, 2007, p. 6)

- Some additional advantages:
  - A way to study problems which are **difficult/impossible to tackle empirically**
  - Provides a (relatively) **quick way to test hypotheses**, make predictions for empirical studies
  - Provides a **baseline** against which to judge alternative explanations

3.2 Example 1: Liljencrants and Lindblom (1972)

- LL are interested in cross-linguistic tendencies in vowel inventories, e.g. (simplified):
  - 3–7 vowel inventories tend to have only front & back vowels ({i, a, u}, {i, e, a, u}...)
  - Certain regions of formant space preferred
  - 5–7 vowel inventories most common
- Investigates to what extent these can be explained by maximizing *perceptual contrast*
- Model:[3] n vowels in $F1, F2'$ space, distance r_{ij} between vowels i and j
  - Choose positions which minimize 'energy': $E = \sum_{i=1}^{n-1} \sum_{j=0}^{i-1} 1/r_{ij}^2$
  - Intuitively, maximize perceptual distance among all vowels in inventory
- Results: Fig. 1

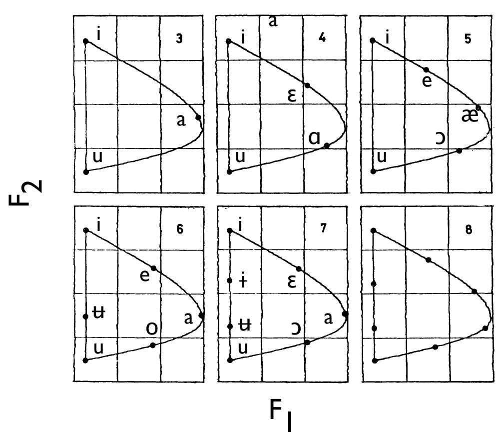


Figure 1: Adapted from Liljencrants and Lindblom (1972).

  - For n =3–5, optimal system ≈ same as most common system cross-linguistically
  - Certain regions of formant space preferred (e.g. point vowels)
  - Central vowels only introduced for $n > 9$
- What can we learn from this model?
  - *Existence proof*: Some major trends in V inventories *could* emerge from just maximizing dispersion. (We'll see in Week 2: can get similar results from a model with no explicit optimization.)

[3]Note that this is not a model of sound change, per se.

- *Communication/explicitness*: LL model is simple, explicit; has served as a baseline for much subsequent work on how V inventory tendencies could emerge from functional constraints.
- *Counterintuitive results*: A preference for peripheral vowels at low n emerges, while other tendencies do not (at what n central vowels become common, which central vowels).
- *Qualitative predictions*: Every vowel inventory should have point vowels, central vowels $\Rightarrow$ front and back vowels, central vowels become more likely for higher n.
- *Baseline*: Maximizing perceptual contrast alone is enough to get some major trends in V inventories, but not many others.

3.3 Example 2: Baker (2008)

- Baker is interested in why universal phonetic pressures don't *inevitably* lead to change
- Investigates one particular hypothesis: change is the accumulation of error (essentially the Neogrammarian perspective)
  - Speakers are connected in a social network based on relative prestige (Fig.2, left)
  - Agents 'speak' to other agents they know, and update their internal representations accordingly
  - Productions are subject to a small amount of random noise

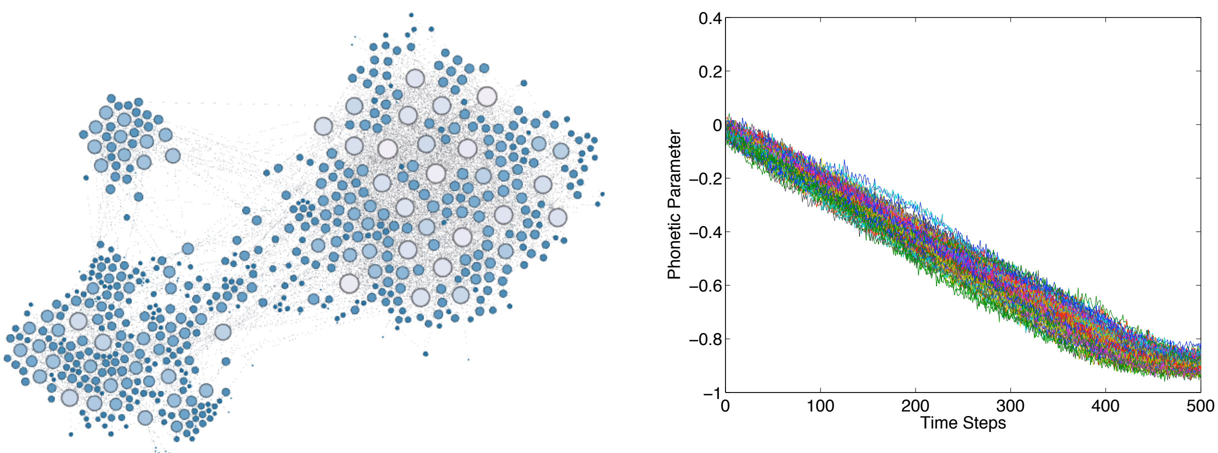


Figure 2: Left: An example of a social network with edges based on social prestige. Right: Invariable sound change produced by a phonologization of coarticulation model (from Baker, 2008).

- **What can we learn from this model?**
  - *(Non-)existence proof*: S-shaped change doesn't emerge from this simple type of model: if a fixed lenition bias is added, change is inevitable, apparently regardless of parameter settings (Fig. 2, right)
  - *Explicitness:* lots of moving parts that need values.
  - *Counterintuitive results:* accumulation-of-error leading to change has struck many researchers as intuitive, but at least in some simple form, it doesn't appear to be tenable.
  - *Predictions:* differences in social structure will impact rates, but not qualitative outcomes?
  - *Baseline:* accumulation-of-error on its own is apparently not enough to get observed change.

4 How do we model?

4.1 General requirements

1. *Knowledge state of individuals*: What information do individuals store? (discrete/continuous; dimensionality)
2. *Network/social structure*: who is data transmitted to/from?
3. *Assumptions about communication* (perception, production): how is data communicated?
4. *Learning algorithm*: how does an individual's state change as a result of receiving data?

4.2 Implementational approaches

- **Analytic** (McElreath and Boyd, 2007)
  - Tries to understand the consequences of a (typically) small number of assumptions in a statistical/mathematical framework (subsumes approaches such as regression, etc.).
  - In literature on change, *dynamical systems* analyses are common:
    1. Specify the distribution of parameter(s) of interest, often expressed as a proportion of members of the population that display a given property.
    2. Compute the change in proportion/frequency after one timestep.
    3. Explore long-term consequences explicitly (if possible) or via equilibrium analysis, etc.
  - Sound change Ex: Niyogi (2006); Sonderegger and Niyogi (2010)

- **Simulation-based** (Gilbert and Terna, 2000; Gilbert, 2007):
  - Explore the (usually non-linear) interaction of a (potentially) large number of parameters.
  - Popular in social sciences b/c allows modeling large populations of heterogenous individuals.
  - A typical *agent-based simulation* architecture:
    1. Agents are specified for production, learning, and memory strategies.
    2. Determine rules governing agent interaction.
    3. Analysis involves examining both internal state of individual agents as well as the results of collective behavior.
  - Sound change Ex: de Boer (2001); Stanford and Kenny (2013); Kirby (2013); Pierrehumbert (2001); Wedel (2006)

- **Combined approaches**
  - Simulation may also be employed when analysis is intractable, i.e. when continuing to assuming homogenous populations.
  - Sound change Ex: Kirby and Sonderegger (2013)

4.3 Pros and cons

Analytic	Simulation
Possibility of proofs of general case as opposed to existence proofs/reasoning based on finite number of cases ('equations talk').	Increased realism in parameterizing the problem in arbitrary levels of detail (potentially a disadvantage; realism/interpretability tradeoff).
Potentially faster/easier to explore parameter space more fully.	Lets you tackle analytically intractable problems (i.e., almost everything).
Common language means maximal explicitness, ease of communication.	Insights into the emergence of macro level phenomena from micro level actions.
Pressure for simplicity aids interpretability.	Simulation is easy (or at least more accessible)!

5 The next two weeks

- While not vast, literature is sizeable
- Our emphasis will be on depth rather than breadth, and hands-on simulation using a few models:
  - Individual phonetic change Pierrehumbert (2001)
  - Evolution of inventories de Boer (2000)
  - Change in populations Kirby and Sonderegger (2015)

6 Themes and questions

- **Exploring the parameter space**
- Sound change in individuals vs. community
- Modeling sound change vs. modeling language change writ large
- Complexity, tractability, running time
- "A map of the Empire the size of the Empire"

References

Auer, P. and Hinskens, F. (2005). The role of interpersonal accommodation in a theory of language change. In Auer, P., Hinskens, F., and Kerswill, P., editors, *Dialect change: convergence and divergence in European languages*, pages 335–357. Cambridge University Press, Cambridge.

Baker, A. (2008). Addressing the actuation problem with quantitative models of sound change. In *Proceedings of the 31st Annual Penn Linguistics Colloquium*, volume 14(1), pages 29–41.

de Boer, B. (2000). Self-organization in vowel systems. *Journal of Phonetics*, 28(4):441–465.

de Boer, B. (2001). *The origins of vowel systems.* Oxford University Press, Oxford.

Campbell, L. (1998). *Historical Linguistics: an introduction.* MIT Press, Cambridge, MA.

Garrett, A. (2014). Sound change. In Bowern, C. and Evans, B., editors, *The Routledge handbook of historical linguistics*, pages 227–248. Routledge, London and New York.

Gilbert, N. (2007). Computational social science: Agent-based social simulation. In Phan, D. and Amblard, F., editors, *Agent-based modelling and simulation*, pages 115–134. Bardwell, Oxford.

Gilbert, N. and Terna, P. (2000). How to build and use agent-based models in social science. *Mind & Society*, 1(1):57–72.

Harrington, J., Palethorpe, S., and Watson, C. (2000). Monophthongal vowel changes in Received Pronunciation: An acoustic analysis of the Queen's Christmas broadcasts. *Journal of the International Phonetic Association*, 30(1-2):63–78.

Hay, J., Warren, P., and Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*, 34(4):458–484.

Hay, J. B., Pierrehumbert, J. B., Walker, A. J., and LaShell, P. (2015). Tracking word frequency effects through 130 years of sound change. *Cognition*, 139:83–91.

Hock, H. H. (1991). *Principles of historical linguistics.* Mouton de Gruyter, Berlin, 2nd edition.

Kirby, J. (2013). The role of probabilistic enhancement in phonologization. In Yu, A., editor, *Origin of Sound Patterns: Approaches to Phonologization*. Oxford University Press, Oxford.

Kirby, J. and Sonderegger, M. (2013). A model of population dynamics applied to phonetic change. In *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*. 6 pp.

Kirby, J. and Sonderegger, M. (2015). Bias and population structure in the actuation of sound change. *arXiv:1507.04420 [cs.CL]*. arXiv: 1507.04420.

Labov, W. (1994). *Principles of linguistic change. Volume 1: Internal factors.* Blackwell, Oxford.

Labov, W. (2000). *Principles of linguistic change. Volume 2: Social factors.* Blackwell, Oxford.

Labov, W. (2007). Transmission and diffusion. *Language*, pages 344–387.

Liljencrants, J. and Lindblom, B. (1972). Numerical simulation of vowel quality systems: the role of perceptual contrast. *Language*, pages 839–862.

McElreath, R. and Boyd, R. (2007). *Mathematical models of social evolution: A guide for the perplexed.* University of Chicago Press, Chicago.

Niyogi, P. (2006). *The computational nature of language learning and evolution.* MIT Press, Cambridge, MA.

Phillips, B. (1984). Word frequency and the actuation of sound change. *Language*, 60(2):320–342.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In Bybee, J. and Hopper, P., editors, *Frequency and the emergence of linguistic structure*, pages 137–157. John Benjamins, Amsterdam.

Sankoff, G. (2005). Cross-sectional and longitudinal studies in sociolinguistics. In Ammon, U., Dittmar, N., Mattheier, K., and Trudgill, P., editors, *Sociolinguistics: An international handbook of the science of language and society*, volume 2, pages 1003–13. de Gruyter, Berlin.

Sonderegger, M. and Niyogi, P. (2010). Combining data and mathematical models of language change. In *Proceedings of the 48th Annual Meeting of the ACL*, pages 1019–1029, Uppsala, Sweden. Association for Computational Linguistics.

Stanford, J. N. and Kenny, L. A. (2013). Revisiting transmission and diffusion: An agent-based model of vowel chain shifts across large communities. *Language Variation and Change*, 25(02):119–153.

Wang, W. (1969). Competing changes as a cause of residue. *Language*, 45(1):9–25.

Wedel, A. (2004). Category competition drives contrast maintenance within an exemplar-based production/perception loop. In *Proceedings of the 7th Meeting of the ACL Special Interest Group in Computational Phonology: Current Themes in Computational Phonology and Morphology*, pages 1–10. Association for Computational Linguistics.

Wedel, A. (2006). Exemplar models, evolution and language change. *The Linguistic Review*, 23:247–274.

Weinreich, U., Labov, W., and Herzog, M. (1968). Empirical foundations for a theory of language change. In Lehmann, W. and Malkiel, Y., editors, *Directions for historical linguistics*. University of Texas, Austin.