Iterated learning of vowel harmony with Bayesian agents

DGfS AG4: "Learning Meets Acquisition"

Fred Mailhot

Institute of Cognitive Science
Carleton University

March 5, 2009

Carleton UNIVERSITY

Overview

Overview

Diachronic explanation of synchronic patterns

Increasing push away from synchronic, grammar-based[1] explanations, toward diachronic, phonetic accounts of phonological patterns (cf. Blevins 2004, Hale & Reiss 2007, Ohala 1992, 2005 *inter alia*).

1. reduce implausible teleological aspects of phonological explanation (cf. Zipf 1935)
2. Ockham's Razor (i.e. eliminate explanatory redundancy)
3. ease Mother Nature's burden (e.g. how to evolve a complex UG)

Focus on a particular case here: the development/emergence of *vowel harmony*.

[1] Esp. markedness-based

Q: Where does vowel harmony come from?

John Ohala (1994) provided the standard (but not universally accepted) answer:

> *Vowel harmony is a result of the phonologization of vowel-to-vowel coarticulation.*

Is there any way of verifying this type of claim?

Linguistic background in a nutshell

Vowel harmony To a first approximation, vowels in some domain ("the word") agree with respect to some set of phonological features, both as a lexical generalization, and w.r.t. productive alternations (*e.g.* Finnish backness harmony)

Phonologization *A universal phonetic tendency is said to become "phonologized" when language specific reference must be made to it, as in a phonological rule.* (Hyman 1972:170) *[. . .] phonologization, whereby a phonetic process becomes phonological. . .* (Hyman 1975:171) Variation under physical/physiological control comes to be under cognitive control.

Coarticulation The predictable effects on segments of their neighbours in running speech. *e.g.* the relative frontness/backness of [k] in *keep* vs. *coop*. V-to-V coarticulation occurs across intervening consonants (*cf.* Öhman 1966, Magen 1997 *inter alia*), and is perceivable by listeners (Beddor et al. 2002).

Sidebar: criteria for VH?

How do we decide whether a language "has" vowel harmony?

1. lexical statistics
   - (*e.g.* deviance from expected rate of harmony, given vowel inventory)
2. loanword adaptation
   - (*e.g.* "harmonization" of disharmonic borrowings)
3. synchronic alternations
   - (*e.g.* as in Finnish suffixes above)

I will focus mostly on (i) and a little bit on (ii) here.

Diachronic explanation . . .

For a candidate explanation of this type, we need:

1. demonstrable synchronic variation
2. proof that variation is detectable by listeners
3. models of synchronic knowledge and acquisition
4. *a demonstration that the above + sufficient time can bring about the phenomenon under consideration*

Point 4 runs into problems.

A methodological stumbling block

- No obvious way to verify the diachronic aspect of this kind of explanation
- Modelling gives us a “virtual lab” in which to test these theories, with perfect repeatability and tight control over parameters
- Other benefits of modelling: (i) quantitative data allow theory comparison/choice, (ii) implementation forces us to be very precise about our theory, parameters and auxiliary assumptions

Modelling language change

- Computational/mathematical modelling of language change is (with few exceptions *cf.* Klein 1969) a relatively recent development in linguistics ($\sim$10 years).
- Mostly deals with syntactic change, with some work on morphological change. Little work on phonological change until quite recently.
- Modelling strategies can be classified as *analytic* (equation-based) (*cf.* Niyogi & Berwick 1995 *et seq.*, Komarova & Nowak 2003) or *synthetic* (agent-based).

Synthetic models

A.K.A. *agent-based modelling, multi-agent simulation*

- Individuals explicitly modeled, population-level properties emerge from local interactions.
- Agents have perception/comprehension model, production model, and internal state/grammar)
- Internal state changes on basis of agent-to-agent interaction.

Can be further subclassified as *vertical* or *horizontal* according to information flow.

Agent-based simulation

Vertical information flow: ***iterated learning model*** (Kirby 1999, *et seq.*):

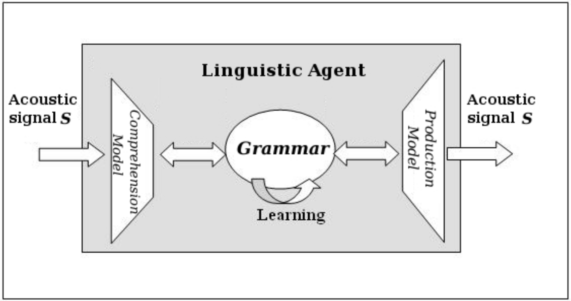
- Typically 2 disjoint subsets of agents, one with fixed internal model ("adults") and one modifiable ("children").[2]
- "Adult" grammars serve as targets. Meant to explicitly capture I-/E-language feedback loop in transmission/acquisition.
- Noisy transmission (and/or information bottleneck) drives change.

My model: simplest possible—one adult, one child per generation

[2] Horizontal info flow: all agents can change internal state, no privileged grammar(s). "Social" vs. "generational" models.

Overview

The agent


Linguistic Agent
Acoustic signal S
Comprehension Model
Grammar
Learning
Production Model
Acoustic signal S

The agent

Mental model: 2 binary features ("high" and "back"). Words are sequences of 4 vowels. Exhaustive (*i.e.* 256-word) lexicon. No morphophonology (yet!).

Production model:

- Continuous articulation: binary features transduced to continuous articulatory parameters; beta-distribution models hyper-/hypoarticulation
- Articulatory synthesizer: equations from de Boer (2000) to generate formant values ($F1, F2$) from articulatory description
- V-to-V coarticulation: contextual variation in $F2$ (approx. analogue of front/back variation)

The agent (cont.)

Comprehension/learning model:

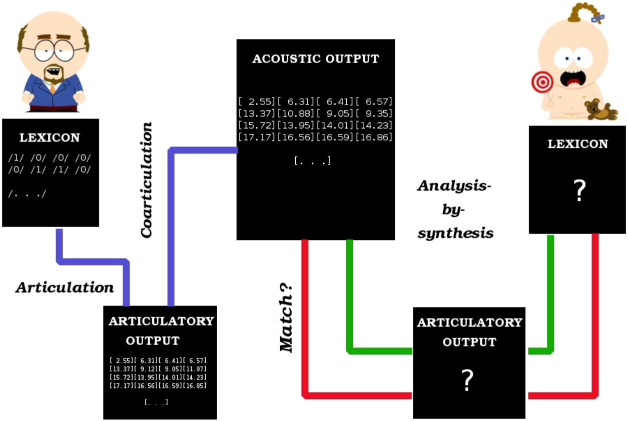
- Acoustic clustering: *k*-means analysis to find acoustic prototypes (GMM-EM too slow, and doesn't reliably find clusters, since data is non-Gaussian)
- Invert articulation-acoustics mapping (this is unrealistic?)
- Feature induction: MAP learning[3] of vowel features from articulatory representations of acoustic prototypes:

$$\hat{h} = \arg\max_{h} \frac{P(D = d|H = h)P(H = h)}{z}$$

- UR induction: Assign URs to words by vector quantization

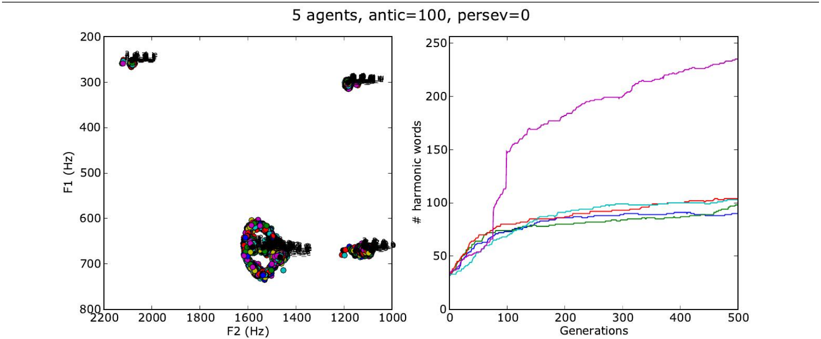
[3] Only uniform priors implemented for now, so this is MLE, strictly speaking.

Information flow

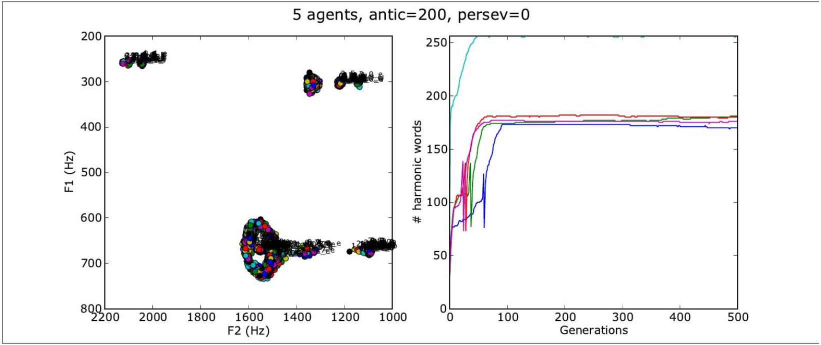


Overview

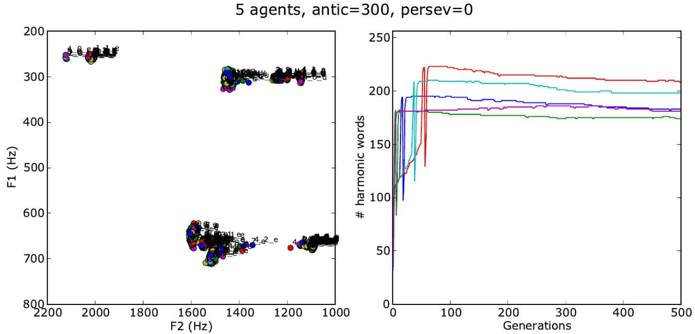
500 generations, antic=100



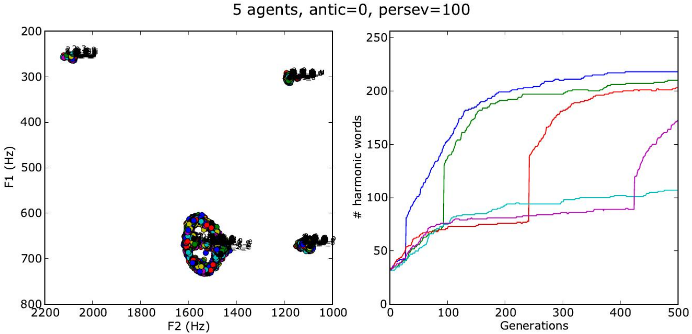
500 generations, antic = 200



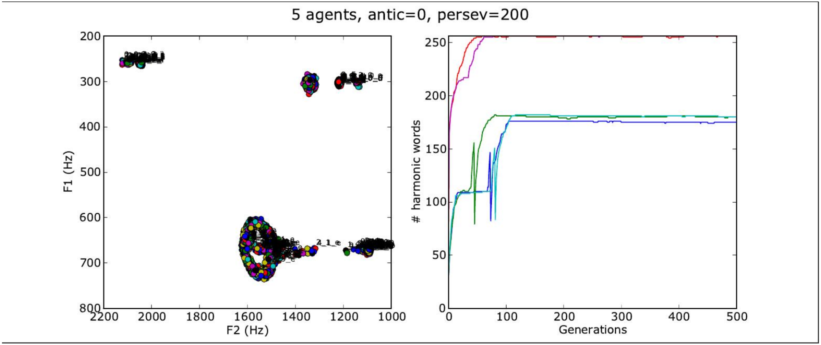
500 generations, antic = 300



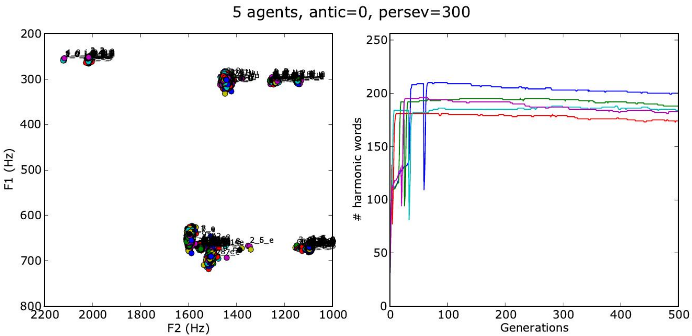
500 generations, persev=100



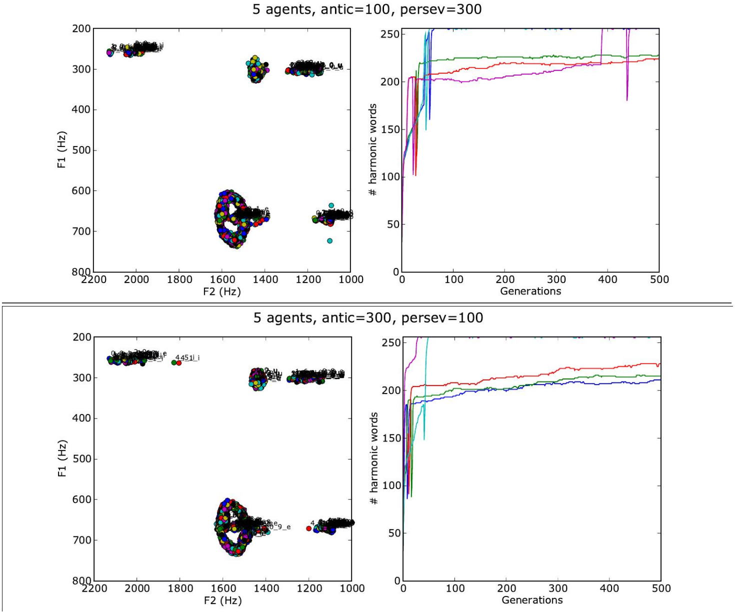
500 generations, persev = 200



500 generations, persev = 300



500 generations, antic & persev



Discussion

- Multiple relatively stable levels of harmony, with different approach speeds depending on amount of coartication.
- Explainable by differential resistance to coarticulation (maybe just different distance in vowel-space for fixed coarticulation)
- Noisy sets show possible partial solution to actuation problem. . . phonologization need not lead inexorably to harmony

Also, informal tests, show that disharmonic loanwords (i.e. acoustic forms corresponding to disharmonic URs) are typically harmonized by agents with harmonic (approx $>$ 200) lexicons.

To do

- MAP learning of words
- Morphophonological alternations
- More realistic data (e.g. stochastic sampling of lexicon at each gen?)
- Social factors (variation in data due to multiple independent input sources)
- Exemplar-based approach—embodying very different assumptions about acquisition and change—is next on the agenda.

Fin

THANK YOU.[4]

[4] Thanks to Ash Asudeh, Jeff Mielke, Lev Blumenfeld and various audiences.

References

Beddor, P. et al. (2002). Language-specific patterns of vowel-to-vowel coarticulation: acoustic structures and their perceptual correlates. *Journal of Phonetics, 30*.

de Boer, B. (2000). *The Origins of Vowel Systems.* OUP.

Blevins, J. (2004.) *Evolutionary Phonology.* OUP.

Hale, M. (2007). *Historical Linguistics: Theory and Method.* Blackwell.

Hare, M., J. Elman. (1995). Learning and morphological change. *Cognition 56*.

Kirby, S. (1999). *Function, Selection and Innateness: The emergence of language universals.* OUP.

Klein, S. (1969). Monte Carlo simulation of language change in Tikopia & Maori. Proceedings of the 1969 conference on Computational linguistics.

Komarova, N., M. Nowak. (2003) Language dynamics in finite populations. Journal of Theoretical Biology, 221. pp. 445-457.

Magen, H. (1997). The extent of vowel-to-vowel coarticulation in English. *Journal of Phonetics,* 25.

Niyogi, P., R. Berwick. (1998). The logical problem of language change: A case study of European Portuguese. *Syntax,1*

Ohala, J. (1992). What's cognitive, what's not, in sound change. In G. Kellermann & M. Morrissey (eds.) *Diachrony within synchrony: Language history and cognition.* Frankfurt/M: Peter Lang Verlag.

Ohala, J. (1993). Coarticulation and Phonology. *Language and Speech, 36*.

Ohala, J. (1994). Towards a universal, phonetically-based, theory of vowel harmony. Third International Conference on Spoken Language Processing. Yokohama, Japan.

Öhman, S. (1966). Coarticulation in VCV utterances: spectrographic measurements. *JASA,* 39.