Iterated learning of vowel harmony with Bayesian agents

DGfS AG4: “Learning Meets Acquisition”

Fred Mailhot

Institute of Cognitive Science
Carleton University

March 5, 2009
Overview

1. Motivation & Background
   - Explanation in phonology
   - VH, Phonologization & Coarticulation
   - Modelling language change

2. Modelling the emergence of harmony
   - The agent
   - The algorithm

3. Results & Discussion
   - Next steps
Overview

1. Motivation & Background
   - Explanation in phonology
   - VH, Phonologization & Coarticulation
   - Modelling language change

2. Modelling the emergence of harmony
   - The agent
   - The algorithm

3. Results & Discussion
   - Next steps

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?
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.
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
Computational/mathematical modelling of language change is (with few exceptions cf. Klein 1969) a relatively recent development in linguistics (~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.
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”).
- “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

---

2Horizontal info flow: all agents can change internal state, no privileged grammar(s). “Social” vs. “generational” models.
Overview

1 Motivation & Background
   - Explanation in phonology
   - VH, Phonologization & Coarticulation
   - Modelling language change

2 Modelling the emergence of harmony
   - The agent
   - The algorithm

3 Results & Discussion
   - Next steps
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 ($F_1$, $F_2$) from articulatory description
- V-to-V coarticulation: contextual variation in $F_2$ (approx. analogue of front/back variation)
**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.
Motivation & Background
Modelling the emergence of harmony
Results & Discussion

The agent
The algorithm

Information flow

LEXICON
/
/
/
/
/
/
/
/
/
/
/
/
/
/
/
/
/
/
/

Articulation

Coarticulation

LEXICON

ACOUSTIC OUTPUT

\[
\begin{array}{c}
2.55 & 6.31 & 6.41 & 6.57 \\
13.37 & 10.88 & 9.95 & 9.95 \\
15.72 & 13.55 & 14.01 & 14.23 \\
17.17 & 16.56 & 16.59 & 16.66 \\
\end{array}
\]

analysis-
by-
synthesis

Match?

LEXICON

ARTICULATORY OUTPUT

\[
\begin{array}{c}
2.55 & 6.31 & 6.41 & 6.57 \\
13.37 & 10.88 & 9.95 & 9.95 \\
15.72 & 13.55 & 14.01 & 14.23 \\
17.17 & 16.56 & 16.59 & 16.66 \\
\end{array}
\]

Iterated learning of vowel harmony
Overview

1. Motivation & Background
   - Explanation in phonology
   - VH, Phonologization & Coarticulation
   - Modelling language change

2. Modelling the emergence of harmony
   - The agent
   - The algorithm

3. Results & Discussion
   - Next steps
500 generations, antic=100

5 agents, antic=100, persev=0
500 generations, antic = 200

5 agents, antic=200, persev=0

F1 (Hz)

F2 (Hz)

Generations

# harmonic words
500 generations, antic = 300

5 agents, antic=300, persev=0
500 generations, persever=100

5 agents, antic=0, persever=100
500 generations, persev = 200

5 agents, antic=0, persev=200

F1 (Hz) vs F2 (Hz)

# harmonic words vs Generations
500 generations, persever = 300

5 agents, anti=0, persever=300
500 generations, antic & persever
Multiple relatively stable levels of harmony, with different approach speeds depending on amount of coarticulation.

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.
THANK YOU.\textsuperscript{4}

\textsuperscript{4}Thanks to Ash Asudeh, Jeff Mielke, Lev Blumenfeld and various audiences.


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


