From Sound Change to Grammar Change: Words, Lexicons, and Learners

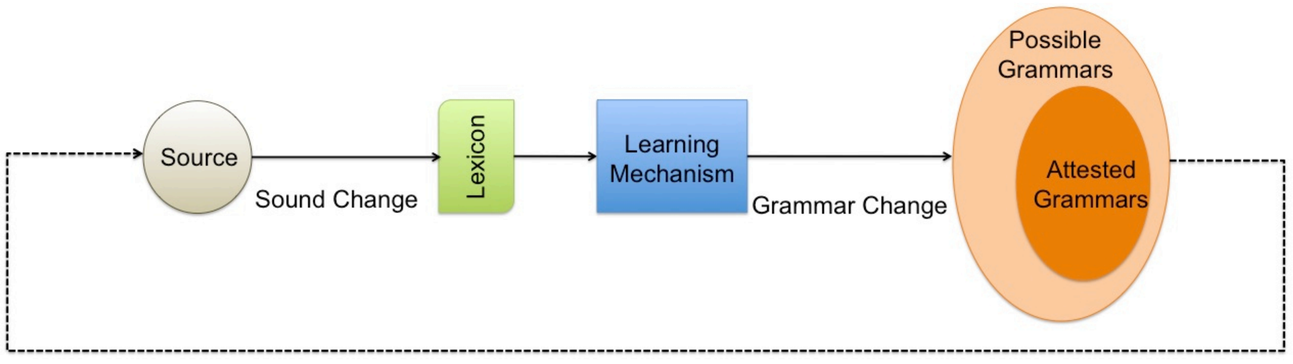
Rebecca Morley
Address correspondence to: morley@ling.osu.edu
THE OHIO STATE UNIVERSITY

Possible/Impossible/Likely sound changes, sequences of changes?

Token frequencies, type distributions, lexicon shape?

Tracking conditional probabilities, assessing significance, categorization of tokens?

Rules/constraints/gradient phonotactics?



The Basic Model

Research Questions

- How does a new phoneme contrast develop over time?
- How likely are unnatural phonotactics to emerge?
- How likely are 'anti-markedness' phonotactics to emerge?

Source: Velar Palatalization

Pseudo-Russian

0. Phonetic Palatalization
   Articulatory Ease
   dakɪ →dak$^{(j)}$ɪ (PIR)

1. Birth of a (Partial) Contrast
   Independent Sound Change:
   Deletion of final jers
   dak$^{(j)}$ɪ > dak^{j}
   dakʊ > dak

2. Expansion of New Phoneme
   Dependent Sound Change:
   re-categorization/re-analysis
   tak$^{(j)}$i > tak^{j}i
   k$^{(j)}$em > k^{j}em
   strak$^{(j)}$ɪlit > strak^{j}ɪlit

3. Expansion of New Contrast
   Dependent Sound Change:
   Contrast Maintenance + Re-analysis
   k^{j}ek > k^{j}eok > k$^{j(e)}$ok >k^{j}ok

Lexicon

Inventory	Segments
• Vowels: {i,ɪ,e,o,a,ʊ,u} • Consonants: {i,ɪ,e,o,a,ʊ,u,b,p,g,k,d,t,m,n,z,s,C$_{1}$,$_{0}$} • Syllables: {70 CV; 70 CVC}	• selected at random until entire set of syllables created **Syllables** • selected at random until entire set of words created

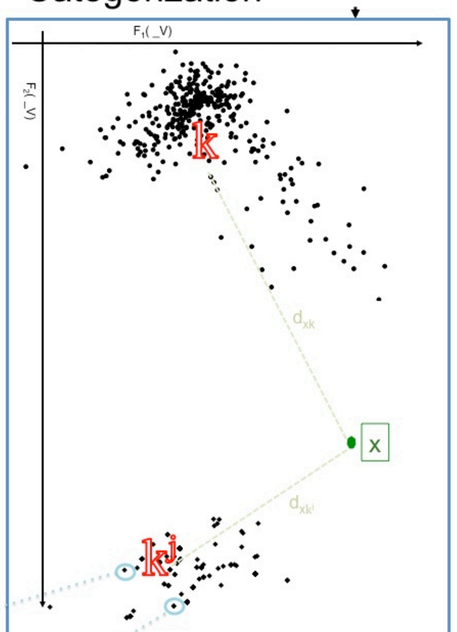
6000 words: [roughly following CELEX(1993) percentages]
- 1-syll words: 1080
- 2-syll words: 2760
- 3-syll words: 2160

Run # 87: …C*b… :

1	ClbmClb	16	ClbmutClb	31	gigslpClb
2	nusClb	17	ClbklClb	32	tusiClb
3	ClbUb	18	ClbkamClb	33	plbClbke
4	ClbUb	19	niClbim	34	deClbag
5	tUtClb	20	motdetClb	35	gezClbdu
6	Clbtu	21	popClbsus	36	ClbblmonClb
7	Clbd	22	tidglmClb	37	blzgozClb
8	zoClb	23	zubuClb	38	mltClbta
9	Clbme	24	ClbpUpi	39	konClbpa
10	Clbo	25	doClbz	40	nUClbtUt
11	Clbdi	26	teClbu	41	bimClbsem
12	mutClb	27	tUbsudClb	42	paClbg
13	mlClb	28	diClbdeg	43	tlpClbgig
14	Clbt	29	detClbdeg	44	bUClbi
15	piClb	30	niClbu		

Learning Mechanisms

Perceptual Categorization



Probability of Re-Categorization
{i,ɪ,e} only:
Bidirectional, but asymmetric
[Guion 1998]

$$P(k>C)_x = e^{-N(\frac{d_{xC}}{d_{xk}})\sqrt{\frac{S_C}{S_k}}} + P_{Phono}(k>C)_x$$

$$P(C>k)_x = Ae^{-N(\frac{d_{xk}}{d_{xC}})\sqrt{\frac{S_k}{S_C}}} + P_{Phono}(C>k)_x$$

A = .25
N = normalization factor [-*ln*(.5)]

* No perceptual bias for ki > Ci over kɪ > Cɪ *

Assessment of Significance

Strength of association:
- Compare to max predictability
- Degree of Implication [*IS*]: $\frac{count(C_1V_2)}{count(V_2)}$

Assessment of Predictability

Predictability Statistic:
[Maddieson & Precoda (1992)]

q ∈ {C,k} [C= k^{j}]
y ∈ {i,ɪ,e,o,a,ʊ,u,b,p,g,k,d,t,m,n,z,s,C$_{1}$,$_{0}$}

Measure of Predictability:

$$Pd_{vw}(y) = P(y_v \mid q_w) - p(y_v)$$

$$= \frac{P(y_v,q_w)}{P(q_w)} - p(y_v)$$

$$E(Pd_{vw}(y)) = \frac{count(y_v q_w)}{count(q_w)} - \frac{count(y_v)}{n}$$

Associations [q=C]
- •Pd$_{12}$(a): $_{1}$ba**C**im$_{0}$
- •Pd$_{21}$(e): $_{1}$**C**ekma$_{0}$
- •Pd$_{13}$(o): $_{1}$tos**C**$_{0}$
- •Pd$_{31}$($_{0}$): $_{1}$zʊ**C**ɪ$_{0}$

Grammatical Bias

Probability of recategorization:

Bidirectional:
k > C {q = C}
C > k {q = k}

$P_{Phono}(k>C)$ = $w(Pd_{12}+Pd_{21}+Pd_{13}+Pd_{31})$

w = .08

Results: Grammars

Simulations
Iterations: New randomly generated lexicon
Generations: Evolution of one lexicon

Each stage computes over entire lexicon (word-by-word)

For Iteration 1:I
1. Create new lexicon

Generation 1:
1. Final deletion: p=.75
2. Re-categorization
   k>C
   C>k
3. Dissimilation: p = .05
4. Vowel shift: p =.70

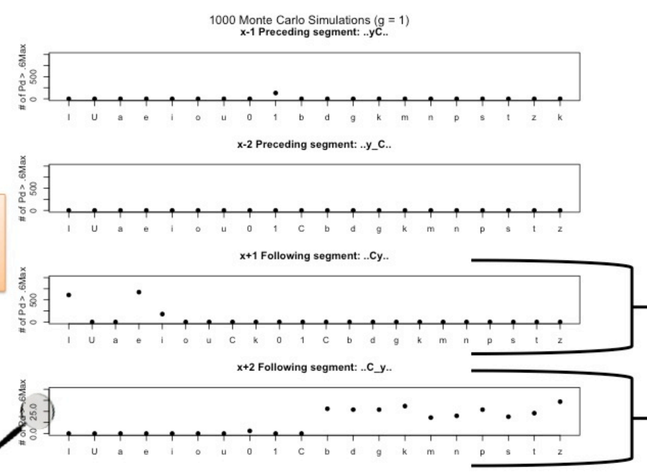
For each word change
- Update phonotactics
- Update category location
- Update category size

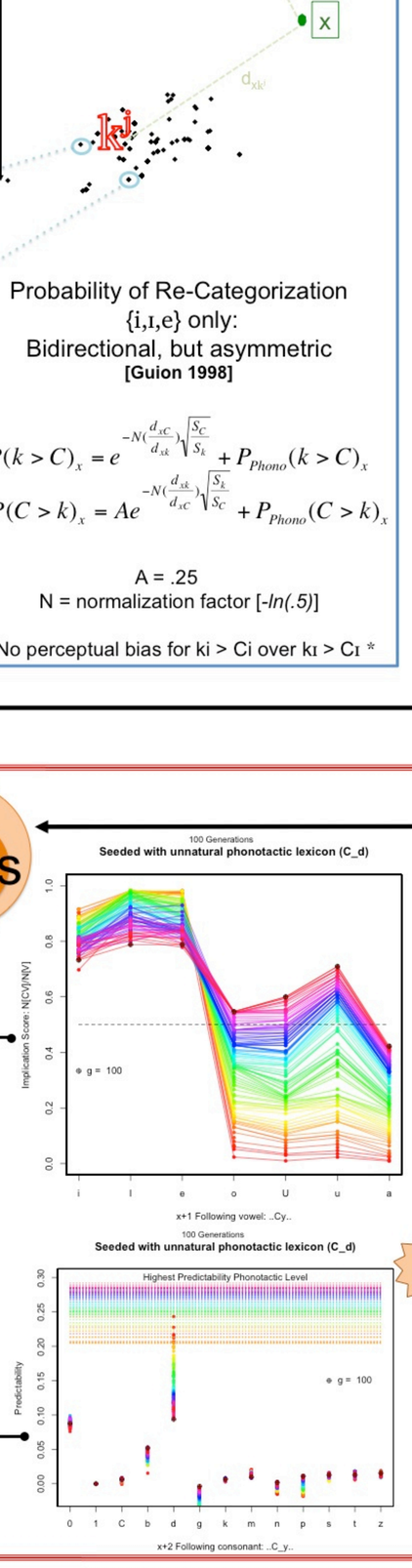
For Generation 2:G
1. Re-categorization
   k>C
   C>k
2. Update lexicon

Loop

Loop

1000 Iterations
1 Generation each





Violations of Universal Implicational Hierarchy
*ki » *kɪ » *ke » *kæ » *ka

100%

- Gradient Violations
  g=1:
  count(Ce)/count(e) = .94
  count(Cɪ)/count(ɪ) = .89
  count(Ci)/count(i) = .70
- Equalization over time
  g=100:
  count(Ce)/count(e) = .79
  count(Cɪ)/count(ɪ) = .71
  count(Ci)/count(i) = .73

3%

Unnatural Phonotactics

- Always co-exist with natural phonotactics
  g=1:
  $Pd_{21}(ɪ)$ = .23
  $Pd_{21}(e)$ = .22
  $Pd_{21}(i)$ = .18
  $Pd_{31}(d)$ = .24
- Persist, but weaken over time

Natural Allophony
Partial Contrast
IS(back vowel) > .4
Full Contrast

- Erosion of Allophony over time
  Bi-modal Distribution:
  'More Likely' with C:{i,ɪ,e}
  'Less Likely' with C:{o,a,ʊ,u}

Conclusions

Anti-markedness Systems:
- ~100% Gradient Violations
  but
- No misperception bias due to vowel quality
- Category seeded at [ɪ] (deletion source)

Unpredictability (contrast) more likely than randomly arising spurious predictability (unnatural patterns).

→Unnatural:	→Natural:
• Randomness • Phonotactics	• Phonotactics • Sound Change • Similarity

Acknowledgments

This work was supported by a Targeted Investment in Excellence grant from the OSU College of Humanities.
Thank you to Paul Smolensky for his input on earlier versions of this work.
I am also very grateful to Andrea Sims for trying to explain Slavic palatalization to me.

Selected References:

•Guion, S. G. (1998). The role of perception in the sounds change of velar palatalization. Phonetica 55: 18-52.
•Kiparsky, P. (2006). The Amphichronic Program vs. Evolutionary Phonology. Theoretical Linguistics 32: 217-236.
•Maddieson, I. and K. Precoda (1992). Syllable Structure and Phonetic Models. Phonology 9: 45-60.
•Ohala, J. (1993). The phonetics of sound change. In C. Jones (ed. Historical Linguistics. 237-278.
•Pierrehumbert, J. B. (2001). Exemplar dynamics: word frequency, lenition and contrast. In J. Bybee and P. J. Hopper (eds.) Frequency effects and the emergence of linguistic structure: John Benjamins. 137-158.
•Pierrehumbert, J. B. (2003). Probabilistic Phonology: Discrimination and Robustness. In R. Bod, J. Hay and S. Jannedy (eds.) Probabilistic Linguistics. Cambridge: The MIT Press. 177-228.
•Townsend, C. E. and L. A. Janda (1996). Common and Comparative Slavic: Phonology and Inflection, Slavica Publishers, Inc.