Bootstrapping Historical Probabilities

Gašper Beguš
Harvard University
begus@fas.harvard.edu

PhoNE
University of Massachusetts, Amherst
April 8, 2017
Questions

- Problems with typology within CB
- Unattestedness of unnatural alternations has been an argument in favor of the Analytic Bias approach to typology (Kiparsky 2006)
- Kiparsky (2006, 2008) lists several diachronic scenarios — combinations of sound changes — that would yield final voicing, yet FV is never attested
- Against Channel Bias: CB wrongly predicts the typology (Bermúdez-Otero 2006)
- Numerous diachronic scenarios can yield unnatural alternations; yet they are not attested
Previous research

- Typology under CB not satisfactory: Rare alternations are infrequent because they are produced by rare sound changes (Blevins 2004)
- Moreton (2008) attempts to quantify phonetic precursors, but phonetic precursors and sound change do not always align
- The most elaborate attempt thus far in Cathcart (2015) lacks MSCR, fails to discriminate alternations from static phonotactic restrictions, and is computationally too demanding and the model
- Fails to yield directly implementable results
Aims

- Propose a quantified model of typology within Channel Bias
On naturalness

- **Natural**: universal phonetic tendencies
- **Unmotivated**: lack phonetic motivations, but not against UPT
- **Unnatural**: operating in the opposite direction from UPT

**Definition of Universal phonetic tendency (UPT)**

UPTs are phonetic processes motivated by articulatory (or perceptual) mechanisms that passively and universally operate in speech production and are typologically common.
Naturalness

- **Unnatural**: operating against universal phonetic tendency (UPT)
Explanations

- The primary diachronic device for explaining unnaturalness is **HYPERCORRECTION** (Ohala 1981).
- Speakers analyze surface form as assimilation and "undo" it.
- "Telescoping", concatenation of sound changes explains unmotivated processes.
- For **UNNATURAL processes**, we need a special scenario ("Blurring Process").
- Unnatural processes are subject of typological debates.
Explanations

- The primary diachronic device for explaining unnaturalness is **hypercorrection** (Ohala 1981).
- Speakers analyze surface form as assimilation and “undo” it.
- “Telescoping”, concatenation of sound changes explains unmotivated processes.
- For **unnatural** processes, we need a special scenario (“Blurring Process”).
- Unnatural processes are subject of typological debates.
- **Blurring Process** is crucial for deriving typology within **CB**.
Models

- **Blurring Process** (Beguš 2016)
- \( A > B / X = UPT \)
- How do we get \( B > A / X \)?
Blurring Process

**Blurring Process**

- A historical model for explaining unnatural phenomena
  
  (a) A set of segments enters complementary distribution
  
  (b) A sound change occurs that operates on the changed/unchanged subset of those segments
  
  (c) Another sound change occurs that blurs the original complementary distribution
Blurring Process

**Blurring Process**

- A historical model for explaining unnatural phenomena
  
  (a) A set of segments enters complementary distribution
  
  (b) A sound change occurs that operates on the changed/unchanged subset of those segments
  
  (c) Another sound change occurs that blurs the original complementary distribution

**Blurring Cycle**

\[ B > C / -X \]
\[ B > A \]
\[ C > B \]
Blurring Process

- **Blurring Process**
- A historical model for explaining unnatural phenomena
  
  (a) A set of segments enters complementary distribution
  
  (b) A sound change occurs that operates on the changed/unchanged subset of those segments
  
  (c) Another sound change occurs that blurs the original complementary distribution

**Blurring Cycle**

- $B > C / \neg X$
- $B > A$
- $C > B$

**Blurring Chain**

- $B > C / X$
- $C > D$
- $D > A$
Blurring Process

- Post-nasal devoicing (Dickens 1984), Hyman (2001)
- In all twelve cases in which PND is reported as a sound change or synchronic alternation, it arises though a combination of 3 sound changes

<table>
<thead>
<tr>
<th>Blurring Cycle</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>D &gt; Z / [−nas]</td>
<td>Avestan</td>
</tr>
<tr>
<td>d &gt; ō/ [−nas]</td>
<td>band</td>
</tr>
<tr>
<td>D &gt; T</td>
<td>Sogdian</td>
</tr>
<tr>
<td>d &gt; t</td>
<td>βand</td>
</tr>
<tr>
<td>Z &gt; D</td>
<td>Yaghnobi</td>
</tr>
<tr>
<td>ō &gt; d</td>
<td>vant</td>
</tr>
</tbody>
</table>

Yaghnobi vant *dasa
Yaghnobi vant das
Blurring Process

- Post-nasal devoicing (Dickens 1984), Hyman (2001)

In all twelve cases in which PND is reported as a sound change or synchronic alternation, it arises though a combination of 3 sound changes.

<table>
<thead>
<tr>
<th>Blurring Cycle</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>D &gt; Z / [−nas]</td>
<td>Avestan band dasa</td>
</tr>
<tr>
<td>d &gt; Ȝ/ [−nas]</td>
<td>Sogdian  QVERIFY dasa</td>
</tr>
<tr>
<td>D &gt; T</td>
<td>Yaghnobi vant *Ȝasa</td>
</tr>
<tr>
<td>Z &gt; D</td>
<td>Yaghnobi vant das</td>
</tr>
</tbody>
</table>
Blurring Process

- Post-nasal devoicing (Dickens 1984), Hyman (2001)
- In all twelve cases in which PND is reported as a sound change or synchronic alternation, it arises though a combination of 3 sound changes

---

<table>
<thead>
<tr>
<th>Blurring Cycle</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>D &gt; Z / [−nas]</td>
<td>Avestan band dasa</td>
</tr>
<tr>
<td>d &gt; ̄d/[−nas]</td>
<td>Sogdian ̄βand ̄đasa</td>
</tr>
<tr>
<td>D &gt; T</td>
<td>Yaghnobi vant ̄*đasa</td>
</tr>
<tr>
<td>d &gt; t</td>
<td>Yaghnobi vant das</td>
</tr>
<tr>
<td>Z &gt; D</td>
<td></td>
</tr>
</tbody>
</table>
Blurring Process

- Post-nasal devoicing (Dickens 1984), Hyman (2001)
- In all twelve cases in which PND is reported as a sound change or synchronic alternation, it arises though a combination of 3 sound changes

Blurring Cycle

<table>
<thead>
<tr>
<th>Blurring Cycle</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>D &gt; Z / [−nas]</td>
<td>Avestan band dasa</td>
</tr>
<tr>
<td>D &gt; T</td>
<td>Sogdian βand ḍasa</td>
</tr>
<tr>
<td>Z &gt; D</td>
<td>Yaghnobi vant *ḍasa</td>
</tr>
<tr>
<td></td>
<td>Yaghnobi vant das</td>
</tr>
</tbody>
</table>
Post-Nasal Devoicing

- Unnatural alternations and phonotactic restrictions
Post-Nasal Devoicing

- Unnatural alternations and phonotactic restrictions
  - Yaghnobi (Xromov 1972)
  - Tswana and Shekgalagari (Solé et al. 2010)
  - Makhuwa and Bube (Janson 1991/1992, Janssens 1993)
  - Konyagi (Merrill 2015)
  - Sicilian and Calabrian (south Italian dial.) (Rohlfs 1949)
  - Murik, Buginese, and Land Dayak (Austronesian) (Blust 2009)
  - Intervocalic devoicing: Berawan and Kiput (Blust 2005, 2013)
  - Voicing after voiceless obstruents in Tarma Quechua (Adelaar 1977, Beguš and Nazarov 2017)
Minimal number of sound changes

- Unnatural processes always require three operating sound changes
- Sound change defined as a change in a single feature value (Picard 1994)
Minimal number of sound changes

- \( B > A / X \)
Minimal number of sound changes

- $B \rightarrow A / X$
- $B \rightarrow A / X$ by definition impossible

Two sound changes? Let $B$ change to $C$, where $B$ and $C$ differ in one feature, but, to be sure, a different feature from the one that separates $A$ and $B$. From this point, it is impossible for an unnatural sound change to arise without a third sound change: $C$ cannot develop directly to $A$, since the two segments differ in two features: feature $F_1$, which distinguishes $A$ and $B$, and feature $F_2$, which distinguishes $B$ and $C$. By definition, two sound changes are required in order to change two features $\rightarrow$ unnatural processes require altogether at least three sound changes.
Minimal number of sound changes

- B > A / X
- B > A / X by definition impossible
- B > C / X
- C > A
Minimal number of sound changes

- $B > A / X$
- $B > A / X$ by definition impossible
- $B > C / X$
- $C > A$

- A and B differ in one feature only
- Two sound change? Let B change to C, where B and C differ in one feature, but, to be sure, a different feature from the one that separates A and B. From this point, it is impossible for an unnatural sound change to arise without a third sound change: C cannot develop directly to A, since the two segments differ in two features: feature $F_1$, which distinguishes A and B, and feature $F_2$, which distinguishes B and C.

- By definition, two sound changes are required in order to change two features $\rightarrow$ unnatural processes require altogether at least three sound changes.
Minimal number of sound changes

Minimal Sound Change Requirement (MSCR)

Natural processes arise through a single sound change. Minimally two sound changes have to operate in combination for an unmotivated process to arise. Minimally three sound changes have to operate in combination for an unnatural process to arise.

- A scale of decreased probabilities
  \[ P_\chi(\text{natural}) < P_\chi(\text{unmotivated}) < P_\chi(\text{unnatural}) \]
Outline

1. Introduction
2. Historical treatments of unnaturalness
3. Bootstrapping SC
4. Applications
5. Conclusion and future directions
Model

- The model can be quantified
- **Bootstrapping Sound Changes**
Historical Probabilities of Alternations ($P_X(Alt)$)

The probability that an alternation arises based on the number of sound changes required (MSCR) and their respective probabilities/rates.

$$P(Alt) = P(T_1 \cup T_2 \cup T_3 \cup ... \cup T_n)$$

- We identify probabilities of each sound changes and then calculate joint probability of alternations.
Poisson Stochastic Process


\[ P(T_1) = \int_0^t f_1 t_1 dt_1 \times \int_{t_1}^t f_2 t_2 dt_2 \times \cdots \times \int_{t_{n-1}}^t f_n t_n dt_n \]

where

\[ f_i = \lambda_i e^{-\lambda_i t} \]
Probability of a sound change

- We can disregard temporal dimension and model sound change probabilities from a given sample as a binary outcome model if we stay within the sample.

- Probability of a sound change $A_1$

$$P_{\chi}(A_1) = \frac{\text{number of languages with sound change } A_1}{\text{number of languages surveyed}}$$

- Easier to implement model
Historical Probabilities

- **Historical probabilities**: Bootstrapping Sound Changes (BSC)

  - Joint probabilities of \( n \) number of sound changes
  - If an alternation \( A_x \) requires \( n > 1 \) sound changes to arise \((n \geq 2\) for unmotivated, \( n \geq 3 \) for unnatural),

  \[
P(T_1) = \frac{P(A_1)P(A_2) \cdots P(A_n)}{n!}
\]

  - Binomial sample: \( P_\chi \) is bootstrapped from a product of probabilities based on the number of successes and failures (divided by \( n! \) to account for the ordering of sound changes)
Assumptions

- Assumptions: sample is representative and well-balanced
- Occurrence of sound change properly counted: occurrence in the proto-language vs. daughter languages
- Independence of sound change
  - $A_1$ and $A_2$ independent
  - $P(A_1 | \text{phonemic inventory})$
  - In the current state of the field, we lack sufficiently accurate estimates of sound change probabilities to be able to estimate their conditional probabilities
  - For practical purposes, we can disregard the phonemic inventory and generalize $P(A_1)$
  - If two sound changes occur in two related languages, they are considered independent
Outline

1. Introduction
2. Historical treatments of unnaturalness
3. Bootstrapping SC
4. Applications
5. Conclusion and future directions
Historical Probabilities

- What can we do with historical probabilities?
  1. Compare two alternations and perform inferential statistics: is historical probability of alternation X significantly more frequent than Y?
  2. Predict attestedness/unattestedness in a given sample
  3. Go below “zero probabilities”
  4. Identify historically equiprobable processes
PND vs. FV

- Post-nasal devoicing (PND) and final voicing (FV)
- Kümmel (2007) surveys approximately 200 languages
- Blurring cycle for PND:
  - 56 languages fricativization of voiced stops
  - 11 languages unconditioned devoicing
  - 37 languages unconditioned occlusion
- Final voicing (FV) (from Kiparsky 2006):
  - The only scenario in Kiparsky (2006) that results in alternation and involves less than four sound changes
    - 10 languages degemination
    - 45 languages post-vocalic voicing
    - 3 languages final degemination
Historical Probabilities

- Estimates of $P_\chi$ were bootstrapped using the boot package (Canty and Ripley 2016, Davison and Hinkley 1997) in R (R Core Team). The following 95% adjusted bootstrap percentile ($BC_a$) intervals were calculated for $P_\chi$(PND) and $P_\chi$(FV).

- **Bootstrapped historical probabilities**
  
  \[ P_\chi(\text{PND}) = [0.02\%, 0.11\%] \]
  \[ P_\chi(\text{FV}) = [0.00\%, 0.02\%] \]
Comparison of $P_\chi$

- We can test whether FV is significantly different from PND
- $P_\chi(\text{PND}) - P_\chi(\text{FV}) = [0.02\%, 0.11\%]$
Comparison of $P_x$

- We can test whether FV is significantly different from PND
- $P_x(\text{PND}) - P_x(\text{FV}) = [0.02\%, 0.11\%]$ 
- Yes, $P_x(\text{PND})$ is significantly higher than $P_x(\text{FV})$. 

Gašper Beguš  
Harvard University — begus@fas.harvard.edu
Comparison of $P_\chi$
Prediction of attestedness

- Do we predict PND to be attested?
- Yes, difference between $P_{\chi}(\text{PND})$ and $P(1/200)$ is $[-0.05\%, 2.47\%]$ 
- But: $P_{\chi}(\text{FV})$ and $P(1/200)$ is $[-0.003\%, 2.86\%]$ 
- $P_{\chi}(\text{FV})$ is not significantly smaller than probability of being attested once in the sample, but $[-0.003\%, 2.86\%]$ and it is significantly different from PND
Going “below zero”

- Why not just calculate probabilities of observed sound changes?
- BSC “reaches below zero”
- Can differentiate unattested processes
Going “below zero”

- FV and intervocalic devoicing both unattested (cf. Yu 2004, recently Sula in Bloyd 2017)
- According to BSC, FV is predicted to be significantly less probable than IVD, but the difference is very small [0.001%, 0.016%]
Going “below zero”
Equiprobable processes

- **Duplication problem:**

- **Analytic Bias** approach (AB): biases in learning affect the typology

- **Channel Bias** approach (CB) assumes phonetic precursors and transmission of language affect the typology
  - Empirical evidence in favor of both
  - Processes that are more difficult to learn are also typologically more rare
  - Processes that are the result of phonologized phonetically motivated sound changes are also typologically frequent

- **Identify equiprobable alternations** for experimental testing
Equiprobable processes

- If typologically rare processes are more difficult to learn and there exist no sound changes that produce them, how can we decide which of the two factors influence the typology?

- Test learnability of processes that are equiprobable historically.

- If an alternation $A_1$ is typologically more common than $A_2$, but $P_\chi(A_1)$ and $P_\chi(A_2)$ are not significantly different, we can test the learnability of $A_1$ and $A_2$. If biases in learnability are confirmed, we would get a direct confirmation of AB without the “duplication” problem outlined above.
Conclusion

- A new quantifiable model of typology within CB
  - Historical probability of FV is small enough for the process not to be attested in a given sample, historical probability PND is significantly higher
  - Based on BSC, Channel Bias would thus correctly predict the observed typology, even in a case that has traditionally been used against CB.

- Comparison of historical probabilities of alternations with inference
- Prediction of attestedness
- Quantification “below zero”
- Identification of historically equiprobable alternations
Future directions

- Phonotactics — MSCR limited to alternations, but unnatural phonotactics result from “blurring process” as well
  (Beguš and Nazarov 2017)

- Further predictions: unmotivated alternations involving multiple features predicted to be less frequent compared to those requiring less features

- Larger, balanced samples

- $P_\chi$ implementable to MaxEnt models of typology (Wilson 2008, White 2017)
  - *Division of Labor Hypothesis*
    - Analytical bias determines $\sigma^2$, channel bias determines $\mu$

- Artificial grammar learning experiments

- Model AB and CB influences on typology
References


———. 2015. Intervocalic Devoicing in Kiput and Berawan Dialects. Presentation at the 22nd AFLA, McGill University.


References


References


References


References


Thank you!