Unnatural Trends in the Lexicon: Diachrony and Synchrony

Gašper Beguš

Harvard University
begus@fas.harvard.edu

Phonology Circle
Massachusetts Institute of Technology
November 7, 2016
Gradient phonotactics


- Little has been said about intersection of the two: unnatural gradient phonotactics
Gradient phonotactics


- Little has been said about intersection of the two: unnatural gradient phonotactics

- Question: Can gradience be unnatural?
Gradient phonotactics


- Little has been said about intersection of the two: unnatural gradient phonotactics

- **Question**: Can gradience be unnatural?

- Trends in the lexicon that operate in an unnatural direction

- Two cases:
  - Tarma Quechua stop voicing
  - Berawan dialects stop devoicing
Naturalness

- What is “unnatural”?

- **Natural**: phonetically motivated
- **Unmotivated**: lack phonetic motivations
- **Unnatural**: operating in the opposite direction from universal phonetic tendency

- UPT: phonetic processes motivated by articulatory (or perceptual) mechanisms that passively and universally operate in speech production and are typologically common.
Introduction

Phonological typology

Two approaches:

- **Analytic bias**
- **Channel bias**

(Moreton 2008)
Introduction

- **Phonological typology**
  - Two approaches: 
    - **Analytic bias**
    - **Channel bias**
  - Literature often treats them as mutually exclusive

(Moreton 2008)
Introduction

- **Phonological typology**
  - Two approaches:  
    - **Analytic bias**
    - **Channel bias**
  - Literature often treats them as mutually exclusive
  - *But*: Both play a role?
  - The goal of phonological theory: model influences of AB and CB

(Moreton 2008)
Introduction

Many open questions:

- Are natural alternations possible?
- Restricted \( \text{CON} \): some constraints not learnable?
- How unnatural processes arise?
- Is sound change natural?
Outline

1. Introduction
2. Data
3. Background
4. Diachrony
5. Typology
6. Theoretical implications
7. Conclusion
Outline

1. Introduction
2. Data
3. Background
4. Diachrony
5. Typology
6. Theoretical implications
7. Conclusion
Outline

1. Introduction
2. Data
3. Background
4. Diachrony
5. Typology
6. Theoretical implications
7. Conclusion
Outline

1. Introduction
2. Data
3. Background
4. Diachrony
5. Typology
6. Theoretical implications
7. Conclusion
Outline

1. Introduction
2. Data
3. Background
4. Diachrony
5. Typology
6. Theoretical implications
7. Conclusion
### Outline

1. Introduction
2. Data
3. Background
4. Diachrony
5. Typology
6. Theoretical implications
7. Conclusion
Outline

1 Introduction
2 Data
3 Background
4 Diachrony
5 Typology
6 Theoretical implications
7 Conclusion
Tarma Quechua, a dialect of Quechua spoken in Tarma, Junín, Peru (Adelaar 1977, Puente Baldoceda 1977, Nazarov 2008)

- Distribution of [voice] in [DOR] and [LAB] stops
- [+voice]: intervocalically, post-consonantally, but not post-nasally
Data

- From Adelaar 1977 and Nazarov 2008

  \[ b, g / C___; C \neq N \]
  \[ b, g / V____V \]
  \[ p, k / \text{elsewhere} \]

- Examples:
- Lexical variation

| ___   | [pirwa]     |
| _____ | [rikra]     |
| R, T   | [wampu]     |
| N___   | [kuba]      |
| V___V  | [takba]     |
Data

- From Adelaar 1977 and Nazarov 2008

\[
\begin{align*}
    & b, g / C_{\_\_}; \; C \neq N \\
    & b, g / V_{\_\_}V \\
    & p, k / \text{elsewhere}
\end{align*}
\]

- Examples:

- Lexical variation

\[
\begin{align*}
    & \#_{\_\_} \quad \text{[pirwa]} \\
    & \_R, T \quad \text{[rikra]} \\
    & N_{\_\_} \quad \text{[wampu]} \\
    & V_{\_\_}V \quad \text{[kuba]} \\
    & R, T_{\_\_} \quad \text{[takba]}
\end{align*}
\]
Data

- Distribution of voicing
Data

- Recount to include non-native vocabulary and statistical tests
- All tokens with [DOR] or [LAB] in TQ vocabulary (Adelaar 1977)
- 1119 tokens: 910 in native vocabulary, 289 in loans from Spanish
Data

Tarma Native

variable

- Voiced
- Voiceless

<table>
<thead>
<tr>
<th></th>
<th>N_</th>
<th>V-V</th>
<th>R_</th>
<th>T_</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Data

Tarma Native+Loans

variable
- Voiced
- Voiceless

N_ | V_V | R_ | T_
Data

- Logistic regression
- Voice
- IV: context and place of articulation with interactions (native vocabulary)
### Data

<table>
<thead>
<tr>
<th>Term</th>
<th>$\beta$</th>
<th>Std. Err.</th>
<th>$z$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-0.52675</td>
<td>0.17847</td>
<td>-2.952</td>
<td>0.00316 **</td>
</tr>
<tr>
<td>R ? T</td>
<td>0.04956</td>
<td>0.26970</td>
<td>0.184</td>
<td>0.85420</td>
</tr>
<tr>
<td>V ? R,T</td>
<td>-0.76124</td>
<td>0.10024</td>
<td>-7.594</td>
<td>3.10e-14 ***</td>
</tr>
<tr>
<td>N ? V,R,T</td>
<td>-0.94159</td>
<td>0.14648</td>
<td>-6.428</td>
<td>1.29e-10 ***</td>
</tr>
<tr>
<td># ? N,V,R,T</td>
<td>-0.81082</td>
<td>0.09321</td>
<td>-8.699</td>
<td>&lt; 2e-16 ***</td>
</tr>
<tr>
<td>place1</td>
<td>0.28282</td>
<td>0.17847</td>
<td>1.585</td>
<td>0.11302</td>
</tr>
<tr>
<td>position1:place1</td>
<td>0.20879</td>
<td>0.26970</td>
<td>0.774</td>
<td>0.43884</td>
</tr>
<tr>
<td>position2:place1</td>
<td>-0.17623</td>
<td>0.10024</td>
<td>-1.758</td>
<td>0.07874 .</td>
</tr>
<tr>
<td>position3:place1</td>
<td>-0.25670</td>
<td>0.14648</td>
<td>-1.752</td>
<td>0.07970 .</td>
</tr>
<tr>
<td>position4:place1</td>
<td>0.01655</td>
<td>0.09321</td>
<td>0.178</td>
<td>0.85908</td>
</tr>
</tbody>
</table>
### Data

<table>
<thead>
<tr>
<th></th>
<th>$\beta$</th>
<th>Std. Err.</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-0.03436</td>
<td>0.12656</td>
<td>-0.272</td>
<td>0.78599</td>
</tr>
<tr>
<td>R ? T</td>
<td>-0.05657</td>
<td>0.21481</td>
<td>-0.263</td>
<td>0.79230</td>
</tr>
<tr>
<td>V ? R,T</td>
<td>-0.65067</td>
<td>0.08229</td>
<td>-7.907</td>
<td>2.64e-15 ***</td>
</tr>
<tr>
<td>N ? V,R,T</td>
<td>-0.65649</td>
<td>0.08389</td>
<td>-7.825</td>
<td>5.06e-15 ***</td>
</tr>
<tr>
<td># ? N,V,R,T</td>
<td>-0.59902</td>
<td>0.05639</td>
<td>-10.624</td>
<td>&lt; 2e-16 ***</td>
</tr>
<tr>
<td>place1</td>
<td>0.57043</td>
<td>0.11772</td>
<td>4.846</td>
<td>1.26e-06 ***</td>
</tr>
<tr>
<td>loan1</td>
<td>0.41520</td>
<td>0.09255</td>
<td>4.486</td>
<td>7.25e-06 ***</td>
</tr>
<tr>
<td>position1:place1</td>
<td>0.11722</td>
<td>0.21473</td>
<td>0.546</td>
<td>0.58515</td>
</tr>
<tr>
<td>position2:place1</td>
<td>-0.15920</td>
<td>0.08180</td>
<td>-1.946</td>
<td>0.05161 .</td>
</tr>
<tr>
<td>position3:place1</td>
<td>-0.15988</td>
<td>0.08355</td>
<td>-1.913</td>
<td>0.05569 .</td>
</tr>
<tr>
<td>position4:place1</td>
<td>0.14757</td>
<td>0.05525</td>
<td>2.671</td>
<td>0.00757 **</td>
</tr>
</tbody>
</table>
Data

- #__ < N__ < V__V < R,T__
- Post-nasally voiced stops universally preferred
  (Hayes and Stivers 2000)
- After voiceless stops voiced stops universally dispreferred
  (voice disagree)
- Intervocally, voiced stops universally preferred
Data

- #___ < N___ < V__V < R,T___
- Post-nasally voiced stops universally preferred
  (Hayes and Stivers 2000)
- After voiceless stops voiced stops universally dispreferred (voice disagree)
- Intervocalically, voiced stops universally preferred
- Gradience goes in the **opposite direction**!
Data

- More voicing after stops than other consonants (Nazarov 2008)
- Only two items with no voicing: [patpatya], [pikpiʃ]
- Reduplication
  - bakbaŋya, lutbi, matga, mutgi, patga, pukba, putga, takba, tikba, tikbi, utga, utgu, witgu, xukbuŋ
- Loan: kutbi < cotpe
### Data

- Post-consonantally voicing after (Nazarov 2008): t, ţ, ţš, k, s, j, x, l, ŧ, r, j, w

<table>
<thead>
<tr>
<th>Pre-Tarma</th>
<th>Tarma</th>
<th>Pre-Tarma</th>
<th>Tarma</th>
</tr>
</thead>
<tbody>
<tr>
<td>aṭška</td>
<td>aṭšga</td>
<td>maṭška</td>
<td>maṭšga</td>
</tr>
<tr>
<td>aṭšpa</td>
<td>aṭšba</td>
<td>arku</td>
<td>argu</td>
</tr>
<tr>
<td>arku</td>
<td>argu</td>
<td>kuṭška</td>
<td>kuṭšga</td>
</tr>
<tr>
<td>awkis</td>
<td>awgis</td>
<td>luṭpi</td>
<td>luṭbi</td>
</tr>
<tr>
<td>aypa</td>
<td>ayba</td>
<td>pilpa</td>
<td>pilba</td>
</tr>
<tr>
<td>ţḷaspu</td>
<td>ţḷasbu</td>
<td>luxpi</td>
<td>luxbi</td>
</tr>
<tr>
<td>ţḷilpi</td>
<td>ţḷilbi</td>
<td>mutki</td>
<td>mutgi</td>
</tr>
<tr>
<td>takpa</td>
<td>takba</td>
<td>tikpa</td>
<td>tikba</td>
</tr>
<tr>
<td>lutfpi</td>
<td>lutfbi</td>
<td>tikpi</td>
<td>tikbi</td>
</tr>
</tbody>
</table>
Data

- Another locus of unnaturalness: **kb, tb, tg**
- Intervocalic stop-stop cluster: $VC_1C_2V$ when $C_2 = [b]$ or $[g]$

\[
\begin{array}{cccc}
TT & TD & DT & DD \\
17 & 69 & 0 & 0 \\
\end{array}
\]

Native only:

\[
\begin{array}{cccc}
TT & TD & DT & DD \\
11 & 68 & 0 & 0 \\
\end{array}
\]

- Clusters that agree in voice preferred
- Pre-voicing preferred
- Tarma Quechua gradience in the **opposite direction**
Data

TT clusters

Count

TT  TD  DT  DD

Gašper Beguš
Harvard University — begus@fas.harvard.edu

Unnatural Trends in the Lexicon: Diachrony and Synchrony
Data

- Phonetically real
- Traces of productivity
Gašper Beguš
Harvard University — begus@fas.harvard.edu

Unnatural Trends in the Lexicon: Diachrony and Synchrony

[atbi]

![Spectral analysis of atbi]

- **Time (s)**: 749.2 to 749.5
- **Frequency (Hz)**: 0 to 5000

The transcribed sound sample [atbi] is analyzed in detail, showing its temporal and spectral characteristics. This analysis provides insights into the diachronic and synchronic aspects of the lexical item under study.
Unnatural Trends in the Lexicon: Diachrony and Synchrony

[atbi]
[ukba]
[ukba]
[asba]
[asba]
[axbi]
[isbi]
[ampa]
[aŋka]
Tarma

- **Traces of productivity**
- **Loanwords**

<table>
<thead>
<tr>
<th>Spanish</th>
<th>Tarma</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>cuculi</td>
<td>kuguli:</td>
<td>white-winged dove</td>
</tr>
<tr>
<td>cotpe</td>
<td>kutbi</td>
<td>an animal from mount.</td>
</tr>
<tr>
<td>sauco</td>
<td>sawgu</td>
<td>magic tree</td>
</tr>
<tr>
<td>vaca</td>
<td>wa:ga</td>
<td>cow</td>
</tr>
<tr>
<td>taruga</td>
<td>taruka</td>
<td>deer’</td>
</tr>
<tr>
<td>dios se lo pague</td>
<td>yusulpa:ki</td>
<td>thank you</td>
</tr>
</tbody>
</table>
Tarma

- Traces of productivity
- Loanwords

<table>
<thead>
<tr>
<th>Spanish</th>
<th>Tarma</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>cuculi</td>
<td>kuguliː</td>
<td>white-winged dove</td>
</tr>
<tr>
<td>cotpe</td>
<td>kutbi</td>
<td>an animal from mount.</td>
</tr>
<tr>
<td>sauco</td>
<td>sawgu</td>
<td>magic tree</td>
</tr>
<tr>
<td>vaca</td>
<td>waːga</td>
<td>cow</td>
</tr>
<tr>
<td>taruga</td>
<td>taruka</td>
<td>deer’</td>
</tr>
<tr>
<td>dios se lo pague</td>
<td>yusulpaːki</td>
<td>thank you</td>
</tr>
</tbody>
</table>

Note that devoicing was never a sound change in TQ
Tarma

- Morphophonological alternation (Creider 1968:12-13)
  - -ba -pa genitive
  - -bax -pax purposive
  - -bita -pita procedentive
  - -bis -pis ‘even, too’

(1) a. wawxi-gi-ba wayi-n
   ‘the house of your brother’

  b. wayi-n-pa pasa-ʃun
   ‘we’re going to walk by way of his house’

  c. tamya-ya-n nuqa-ntʃik-baq
   ‘it is raining now for us’
Tarma

- Morphophonological alternation (Creider 1968:12-13)
  - ba - pa  genitive
  - bax - pax  purposive
  - bita - pita  procedentive
  - bis - pis  ‘even, too’

(1) a. wawxi-gi-\textit{ba} wayi-n
  ‘the house of your brother’

b. wayi-n-\textit{pa} pasa-\textit{jun}
  ‘we’re going to walk by way of his house’

c. tamya-ya-n nuqa-ntʃik-\textit{baq}
  ‘it is raining now for us’
Tarma

- Morphophonological alternation (Creider 1968:12-13)
  - ba - pa genitive
  - bax - pax purposive
  - bita - pita procedentive
  - bis - pis ‘even, too’

(1) a. wawxi-gi-ba wayi-n
   ‘the house of your brother’
  b. wayi-n-pa pasa-ŋun
   ‘we’re going to walk by way of his house’
  c. tamya-ya-n nuqa-ntʃik-baq
   ‘it is raining now for us’

- Limited to some suffixes: -guna, -bura, -gama, and -gasga
  show no alternation
Interim summary

Tarma Native

variable
- Voiced
- Voiceless

N  V_V  R  T

0%  25%  50%  75%  100%
Interim summary

- Significant unnatural trend in the lexicon
- Unnatural gradient phonotactic restriction
- Traces of productivity in morphophonology
Berawan

- Restriction against [+voice] in V__V
- Significant trend in the lexicon: full contrast in [voice] initially, intervocally the natural element dispreferred
- Voicing intervocally (Westbury and Keating 1986)
Intervocalic Devoicing

- Intervocalic (or postvocalic) voicing is common
  \[ T > D / V\_ (V) \]
- The most frequent type of lenition (alongside spirantization, approximatization and others)
  Kaplan (2010)
- 26 of 153 languages have intervocalic lenition
Intervocalic Devoicing

- Good phonetic rationale
- **Passive voicing**: voicing into closure of intervocalic voiceless stops (e.g. German)
  
  Jessen (1998)

- Articulation
  
  Keating (1986)

- P-map: minimal change to achieve phonotactics: $T \rightarrow D$
  
  Kaplan (2010), Steriade (2010)

- Initial devoicing common (cf. English, Yidin)
Intervocalic Devoicing

- Very common sound change
- Reported in over 40 languages

Kümmel (2007)
Intervocalic Devoicing

- Very common sound change
- Reported in over 40 languages

Kümmel (2007)

Intervocalic voicing is a universal **phonetic tendency**.
Intervocalic Devoicing

- Intervocalic *devoicing* is the opposite process
  \[ D > T / V\_\_ (V) \]

- *Unattested* as a synchronic phonological process
- It would operate against the universal phonetic tendency of voicing intervocalic voiceless stops
- P-map: spirantization is perceptually less salient than devoicing in intervocalic position

Steriade (2001), Kaplan (2010)
Berawan

- Analysis of the lexicon in Burkhardt (2014)
- Counts of occurrence: initial, intervocalic
## Berawan

### Labial

<table>
<thead>
<tr>
<th>dialect</th>
<th>voiceless</th>
<th>voiced</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>V__V</td>
</tr>
<tr>
<td>Batu Belah</td>
<td>52</td>
<td>10</td>
</tr>
<tr>
<td>Long Teru</td>
<td>46</td>
<td>13</td>
</tr>
<tr>
<td>Long Jegan</td>
<td>49</td>
<td>10</td>
</tr>
<tr>
<td>Long Terawan</td>
<td>41</td>
<td>11</td>
</tr>
</tbody>
</table>
Berawan

- Alveolar

<table>
<thead>
<tr>
<th>dialect</th>
<th>voiceless</th>
<th>voiced</th>
</tr>
</thead>
<tbody>
<tr>
<td>#__   V__V</td>
<td>#__   V__V</td>
<td></td>
</tr>
<tr>
<td>Batu Belah</td>
<td>56 32</td>
<td>22 4</td>
</tr>
<tr>
<td>Long Teru</td>
<td>54 31</td>
<td>22 2</td>
</tr>
<tr>
<td>Long Jegan</td>
<td>55 32</td>
<td>22 3</td>
</tr>
<tr>
<td>Long Terawan</td>
<td>60 25</td>
<td>21 3</td>
</tr>
</tbody>
</table>
Berawan

- **Velar**

<table>
<thead>
<tr>
<th>dialect</th>
<th>voiceless</th>
<th>voiced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Batu Belah</td>
<td>43</td>
<td>13</td>
</tr>
<tr>
<td>Long Teru</td>
<td>40</td>
<td>11</td>
</tr>
<tr>
<td>Long Jegan</td>
<td>44</td>
<td>10</td>
</tr>
<tr>
<td>Long Terawan</td>
<td>50</td>
<td>14</td>
</tr>
</tbody>
</table>
Data

Batu Belah

<table>
<thead>
<tr>
<th>Labial</th>
<th>Alveolar</th>
<th>Velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>#_</td>
<td>#_</td>
<td>#_</td>
</tr>
<tr>
<td>V_V</td>
<td>V_V</td>
<td>V_V</td>
</tr>
</tbody>
</table>

variable

- Voiced
- Voiceless

# - 0%
# - 25%
# - 50%
# - 75%
# - 100%
V_V - 25%
V_V - 50%
V_V - 75%
V_V - 100%
Data

Long Teru

Variable

Voiced

Voiceless

#_ V_V

Labial

Alveolar

Velar

0%

25%

50%

75%

100%
Data

Long Jegan

<table>
<thead>
<tr>
<th></th>
<th>Labial</th>
<th>Alveolar</th>
<th>Velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
<td>#_</td>
<td>#_</td>
<td>#_</td>
</tr>
<tr>
<td>Voiceless</td>
<td>100%</td>
<td>50%</td>
<td>100%</td>
</tr>
<tr>
<td>Voiced</td>
<td>0%</td>
<td>25%</td>
<td>0%</td>
</tr>
</tbody>
</table>

variable
- Voiced
- Voiceless
Data

Long Terawan

Labial

<table>
<thead>
<tr>
<th>Voiceless</th>
<th>Voiced</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>50%</td>
</tr>
</tbody>
</table>

Alveolar

<table>
<thead>
<tr>
<th>Voiceless</th>
<th>Voiced</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>25%</td>
</tr>
</tbody>
</table>

Velar

<table>
<thead>
<tr>
<th>Voiceless</th>
<th>Voiced</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>25%</td>
</tr>
</tbody>
</table>

variable

- Voiced
- Voiceless
Data

Berawan

<table>
<thead>
<tr>
<th>Category</th>
<th>BB</th>
<th>LJ</th>
<th>LTn</th>
<th>LTu</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voice</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voiced</td>
<td>25%</td>
<td>25%</td>
<td>25%</td>
<td>25%</td>
</tr>
<tr>
<td>Voiceless</td>
<td>75%</td>
<td>75%</td>
<td>75%</td>
<td>75%</td>
</tr>
</tbody>
</table>

voice
voiced
voiceless
### Data

- **Negative binomial:** voice, context, place
- **count:** context * voice + place * voice
### Data

|                  | Estimate | Std. Error | z value | Pr(>|z|) |
|------------------|----------|------------|---------|----------|
| (Intercept)      | 4.0280   | 0.2368     | 17.01   | 0.0000 *** |
| positionintervoc | -0.5653  | 0.2446     | -2.31   | 0.0208 *  |
| voicevoice       | -0.7524  | 0.3826     | -1.97   | 0.0493 *  |
| placelabial      | -0.4285  | 0.3050     | -1.40   | 0.1601    |
| placevelar       | 0.1805   | 0.2912     | 0.62    | 0.5354    |
| positionV:voice+ | -1.9673  | 0.5371     | -3.66   | 0.0002 *** |
| voice+:labial    | 0.7717   | 0.5010     | 1.54    | 0.1235    |
| voice+:velar     | -1.0127  | 0.5498     | -1.84   | 0.0655 .  |

---

Gašper Beguš
Harvard University — begus@fas.harvard.edu

Unnatural Trends in the Lexicon: Diachrony and Synchrony
## Data

|                         | Estimate | Std. Error | z value | Pr(>|z|) |
|-------------------------|----------|------------|---------|---------|
| (Intercept)             | 3.9302   | 0.2026     | 19.40   | 0.0000  |
| positionintervoc        | -0.4224  | 0.2098     | -2.01   | 0.0440  |
| voicevoice              | -0.8086  | 0.3435     | -2.35   | 0.0186  |
| placelabial             | -0.3969  | 0.2634     | -1.51   | 0.1319  |
| placevelar              | 0.1702   | 0.2483     | 0.69    | 0.4931  |
| positionV:voice+        | -2.2198  | 0.5364     | -4.14   | 0.0000  |
| voice+:labial           | 0.8768   | 0.4481     | 1.96    | 0.0504  |
| voice+:velar            | -0.8692  | 0.5019     | -1.73   | 0.0833  |
## Data

|                          | Estimate | Std. Error | z value | Pr(>|z|) |
|--------------------------|----------|------------|---------|----------|
| (Intercept)              | 3.9962   | 0.2328     | 17.17   | 0.0000   |
| positionintervoc         | -0.5174  | 0.2400     | -2.16   | 0.0311   |
| voicevoice               | -0.7822  | 0.3811     | -2.05   | 0.0401   |
| placelabial              | -0.4553  | 0.3009     | -1.51   | 0.1302   |
| placevelar               | 0.2378   | 0.2849     | 0.83    | 0.4039   |
| positionV:voice+         | -2.1535  | 0.5658     | -3.81   | 0.0001   |
| voice+:labial            | 0.8838   | 0.4988     | 1.77    | 0.0764   |
| voice+:velar             | -1.2470  | 0.5674     | -2.20   | 0.0280   |
Data

- Logistic regression

|                     | Estimate | Std. Error | z value | Pr(>|z|) |
|---------------------|----------|------------|---------|----------|
| (Intercept)         | -0.7246  | 0.1493     | -4.85   | 0.0000   |
| positionintervoc    | -1.8927  | 0.5440     | -3.48   | 0.0005   |
| place1              | -0.2328  | 0.1995     | -1.17   | 0.2431   |
| place2              | 0.8484   | 0.1870     | 4.54    | 0.0000   |
Data

- Productivity: loanwords
- Brunei Malay [pəsɪɡuːpən] > Pre-Berawan *səɡuːpən > BB [səkupən]
- Brunei Malay sigup > Pre-Berawan *sigup > BB [sigup]
Data

Berawan

voice
voiced
voiceless
Data

- Significant unnatural trend in the lexicon
- Unnatural gradient phonotactic restriction against voiced stops intervocalically
Data

- How did Berawan and Tarma Quechua phonotactics arise?
<table>
<thead>
<tr>
<th></th>
<th>Outline</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
</tr>
<tr>
<td>2</td>
<td>Data</td>
</tr>
<tr>
<td>3</td>
<td>Background</td>
</tr>
<tr>
<td>4</td>
<td>Diachrony</td>
</tr>
<tr>
<td>5</td>
<td>Typology</td>
</tr>
<tr>
<td>6</td>
<td>Theoretical implications</td>
</tr>
<tr>
<td>7</td>
<td>Conclusion</td>
</tr>
</tbody>
</table>
Background

- Synchronic alternation: \textit{post-nasal devoicing}
Post-Nasal Devoicing

- Post-nasal voicing (T > D / N__) is well attested both as a synchronic process and as a sound change.
Post-Nasal Devoicing

Post-nasal voicing (T > D / N__) is well attested both as a synchronic process and as a sound change

Gk. /ton topo/ → [tondopo]

(Newton 1972)
Post-Nasal Devoicing

- Post-nasal voicing ($T > D / N_-$) is well attested both as a synchronic process and as a sound change.

  Gk. /ton topo/ → [tondopo]

  (Newton 1972)
Post-Nasal Devoicing

- Post-nasal voicing (T $>$ D / N__) is well attested both as a synchronic process and as a sound change.
  
  Gk. /ton topo/ $\rightarrow$ [tondopo]
  
  (Newton 1972)

- Phonetically well motivated: nasal leak (airflow), volume expansion

- Passive voicing in post-nasal position even for languages like English

- PNV is a **universal phonetic tendency**
  
  (Hayes and Stivers 1995)
Post-Nasal Devoicing

- Clear phonetic reasons for why unnatural post-nasal devoicing is universally avoided
Post-Nasal Devoicing

- Clear phonetic reasons for why unnatural post-nasal devoicing is universally avoided

\[ D > T / N \]
Post-Nasal Devoicing

- Clear phonetic reasons for why unnnatural post-nasal devoicing is universally avoided

\[ D > T / N \]

- As phonological rule attested in Tswana and Shekgalagari

Sh. /χu-m-bón-á/ → [χumpóná]
/χu-m-du3-a/ → [χuntu3a]

(Hyman 2001, Zsiga et al. 2006, Solé et al. 2010)
Post-Nasal Devoicing

- Clear phonetic reasons for why unnatural post-nasal **devoicing** is universally avoided
  
  \[ D > T \text{ / } N \]

- As phonological rule attested in Tswana and Shekgalagari

  Sh. \( /χυ-m-bόn-ά/ \rightarrow [χυmpόnά] \)
  \( /χυ-m-du3-a/ \rightarrow [χυntu3a] \)

Post-Nasal Devoicing

- Post-nasal devoicing as sound change reported for eleven languages and dialects from five language families.
Post-Nasal Devoicing

- Post-nasal devoicing as sound change reported for eleven languages and dialects from five language families
  
  - 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)
If PND as sound change indeed occurred, it would mean that either

(a) sound change *can* operate against the principles of phonetic naturalness or
(b) PND is phonetically well-motivated and natural process.
Explanations of Post-Nasal Devoicing

- Various attempts to explain post-nasal devoicing as a sound change
- Difficulties explaining why a sound change should operate against the strong phonetic tendency
- Limited to only one instance of PND and the data is examined in isolation
- The explanations often involve reconstructing hypothetical and unattested dialects for which there is no comparative evidence
Explanations of Post-Nasal Devoicing

- Dickens (1984) and Hyman (2001) argue for a combination of sound changes.
- In the absence of typological parallels, one might judge an explanation that operates with a single (albeit unnatural) sound change more parsimonious.
- Essentially the correct explanation.
Yaghnobi

- Xromov (1972, 128): $D > T / N$

<table>
<thead>
<tr>
<th>Yaghnobi</th>
<th>Sogdian</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>'γantum</td>
<td>γandum</td>
<td>’wheat’</td>
</tr>
<tr>
<td>ji'kampa</td>
<td>əʃkamb</td>
<td>’stomach’</td>
</tr>
<tr>
<td>sank(a)</td>
<td>sang</td>
<td>’stone’</td>
</tr>
<tr>
<td>ran'kiña</td>
<td>rang</td>
<td>’color’</td>
</tr>
<tr>
<td>un'kuft</td>
<td>anguʃt</td>
<td>’finger’</td>
</tr>
<tr>
<td>'tʃintir</td>
<td>tʃəndər</td>
<td>postp.</td>
</tr>
<tr>
<td>-ant</td>
<td>-and</td>
<td>3rd pl.</td>
</tr>
</tbody>
</table>
Yaghnobi

- Xromov (1972, 128): D > T / N

<table>
<thead>
<tr>
<th>Yaghnobi</th>
<th>Sogdian</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ɣantum</td>
<td>ɣandum</td>
<td>‘wheat’</td>
</tr>
<tr>
<td>ʃi'kampa</td>
<td>ʃkamb</td>
<td>‘stomach’</td>
</tr>
<tr>
<td>sank(a)</td>
<td>sang</td>
<td>‘stone’</td>
</tr>
<tr>
<td>ran'kiːna</td>
<td>rang</td>
<td>‘color’</td>
</tr>
<tr>
<td>un'kuʃt</td>
<td>anguʃt</td>
<td>‘finger’</td>
</tr>
<tr>
<td>'tʃintir</td>
<td>ʃəndər</td>
<td>postp.</td>
</tr>
<tr>
<td>-ant</td>
<td>-and</td>
<td>3rd pl.</td>
</tr>
</tbody>
</table>
Yaghnobi

Proposal:

**Stage 1** Fricativization of voiced stops except after nasals

*(complementary distribution)*

\[ D > \text{D} / [-\text{nas}] \]
Yaghnobi

Proposal:

Stage 1 Fricativization of voiced stops except after nasals
(complementary distribution)

\[ D > D / [-nas] \]

Stage 2 Unconditioned devoicing of voiced stops

\[ D > T \]
Yaghnobi

- Direct historical evidence for such a proposal: Sogdian
Yaghnobi

- Direct historical evidence for such a proposal: **Sogdian**
- **Stage 1** is directly attested:

<table>
<thead>
<tr>
<th>Proto-Iranian</th>
<th>Sogdian</th>
<th>Yaghnobi</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>D &gt; D / [-nas]</td>
<td>βand</td>
</tr>
<tr>
<td></td>
<td>D &gt; T</td>
<td>vant</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Avestan</th>
<th>Sogdian</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>dasa</td>
<td>ḍasa</td>
<td>‘ten’</td>
</tr>
<tr>
<td>gari-</td>
<td>γarí</td>
<td>‘mountain’</td>
</tr>
<tr>
<td>asəŋga-</td>
<td>sang</td>
<td>‘stone’</td>
</tr>
</tbody>
</table>
Yaghnobi

- Sogdian and Yaghnobi offer further crucial insights
- Alveolar series of stops undergo occlusion “back to stops”
  - $d > \delta$ / [-nas]_
  - $d > t$
  - $\delta > d$

<table>
<thead>
<tr>
<th>Sogdian</th>
<th>Yaghnobi</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma d$</td>
<td>$x t$</td>
</tr>
<tr>
<td>$\beta d$</td>
<td>$f t$</td>
</tr>
<tr>
<td>$z d$</td>
<td>$s t$</td>
</tr>
</tbody>
</table>
Tswana, Shekgalagari

- Solé et al. (2010):

<table>
<thead>
<tr>
<th>No N-prefix</th>
<th>N-prefix</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>χu-pak-a</td>
<td>χu-m-pak-a</td>
<td>‘to praise’</td>
</tr>
<tr>
<td>χu-tut-a</td>
<td>χu-n-tut-a</td>
<td>‘to respect’</td>
</tr>
<tr>
<td>χu-cúb-á</td>
<td>χu-ŋ-cúb-á</td>
<td>‘to beat’</td>
</tr>
<tr>
<td>χu-kél-a</td>
<td>χu-ŋ-kél-a</td>
<td>‘to show’</td>
</tr>
<tr>
<td>χu-bón-á</td>
<td>χu-m-pón-á</td>
<td>‘to see’</td>
</tr>
<tr>
<td>χu-duz-a</td>
<td>χu-n-tuz-a</td>
<td>‘to annoint’</td>
</tr>
<tr>
<td>χu-fís-a</td>
<td>χu-ŋ-cís-a</td>
<td>‘to feed’</td>
</tr>
<tr>
<td>χu-at-a</td>
<td>χu-ŋ-gat-a</td>
<td>‘like’</td>
</tr>
<tr>
<td>χu-hịn-a</td>
<td>χu-m-pʰịn-a</td>
<td>‘defeat’</td>
</tr>
</tbody>
</table>
### Tswana, Shekgalagari

- **Solé et al. (2010):**

<table>
<thead>
<tr>
<th>No N-prefix</th>
<th>N-prefix</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>χʊ-pak-a</td>
<td>χʊ-m-pak-a</td>
<td>‘to praise’</td>
</tr>
<tr>
<td>χʊ-tut-a</td>
<td>χʊ-n-tut-a</td>
<td>‘to respect’</td>
</tr>
<tr>
<td>χʊ-cúb-á</td>
<td>χʊ-ŋ-cúb-á</td>
<td>‘to beat’</td>
</tr>
<tr>
<td>χʊ-kɛl-a</td>
<td>χʊ-ŋ-kɛl-a</td>
<td>‘to show’</td>
</tr>
<tr>
<td>χʊ-bón-á</td>
<td>χʊ-m-póń-á</td>
<td>‘to see’</td>
</tr>
<tr>
<td>χʊ-duñ-a</td>
<td>χʊ-n-tuñ-a</td>
<td>‘to anoint’</td>
</tr>
<tr>
<td>χʊ-jís-a</td>
<td>χʊ-ŋ-cís-a</td>
<td>‘to feed’</td>
</tr>
<tr>
<td>χʊ-at-a</td>
<td>χʊ-ŋ-gat-a</td>
<td>‘like’</td>
</tr>
<tr>
<td>χʊ-hın-a</td>
<td>χʊ-m-pʰın-a</td>
<td>‘defeat’</td>
</tr>
</tbody>
</table>
Tswana, Shekgalagari

- Proposal in Dickens (1984) and Hyman (2001):
Tswana, Shekgalagari

- Proposal in Dickens (1984) and Hyman (2001):
- **Complementary distribution**, then unconditioned devoicing.
Tswana, Shekgalagari

- Proposal in Dickens (1984) and Hyman (2001):
- **Complementary distribution**, then unconditioned devoicing.

**Stage 1**  Fricativization of voiced stops except after nasals
  (complementary distribution)

\[ D > \mathcal{D} / [-\text{nas}]\]
Tswana, Shekgalagari

- Proposal in Dickens (1984) and Hyman (2001):
- **Complementary distribution**, then unconditioned devoicing.

**Stage 1**  Fricativization of voiced stops except after nasals  
(complementary distribution)  
\[
\text{D} \quad \text{D} / [-\text{nas}]\]

**Stage 2**  Unconditioned devoicing of voiced stops  
\[
\text{D} \quad \text{T}
\]

**Stage 3**  Unconditioned occlusion of voiced fricatives  
\[
\text{D} \quad \text{D}
\]
Tswana, Shekgalagari

- Dialectal evidence (Solé et al. 2010)
- Tswana has systems with *fricativization except after nasals* and systems with *unconditioned devoicing*
Tswana, Shekgalagari

- Dialectal evidence (Solé et al. 2010)
- Tswana has systems with **fricativization except after nasals** and systems with **unconditioned devoicing**
- Combination of these two dialects → **post-nasal devoicing**

<table>
<thead>
<tr>
<th></th>
<th>#ba</th>
<th>aba</th>
<th>mba</th>
</tr>
</thead>
<tbody>
<tr>
<td>devoicers</td>
<td>#pa</td>
<td>aba</td>
<td>mba</td>
</tr>
<tr>
<td>leniters</td>
<td>#βa</td>
<td>aβa</td>
<td>mba</td>
</tr>
<tr>
<td>post-nasal devoi.</td>
<td>#ba</td>
<td>aba</td>
<td>mpa</td>
</tr>
</tbody>
</table>
Tswana, Shekgalagari

- Dialectal evidence (Solé et al. 2010)
- Tswana has systems with **fricativization except after nasals** and systems with **unconditioned devoicing**
- Combination of these two dialects → post-nasal devoicing

<table>
<thead>
<tr>
<th></th>
<th>#ba</th>
<th>aba</th>
<th>mba</th>
</tr>
</thead>
<tbody>
<tr>
<td>devoicers</td>
<td>#pa</td>
<td>apa</td>
<td>mpa</td>
</tr>
<tr>
<td>leniters</td>
<td>#βa</td>
<td>aβa</td>
<td>mba</td>
</tr>
<tr>
<td>post-nasal devoi.</td>
<td>#ba</td>
<td>aba</td>
<td>mpa</td>
</tr>
</tbody>
</table>

- Crucially, fricatives got occluded back to stops
Konyagi

- Recently PND identified as sound change in Konyagi
- The model exactly predicts the development
- All stops fricativize in Konyagi (dialectal data), even nasals
Konyagi

(Data from Merrill, ms.)

P | p t c k f j x b d j g w y l m n ŋ ĩ ñ
Konyagi

(Data from Merrill, ms.)

\[P \mid p \ t \ c \ k \ f \ j \ x \ b \ d \ j \ g \ w \ y \ l \ m \ n \ \tilde{n} \ \eta\]

\[P \mid f \ \tilde{r} \ \tilde{x} \ f \ j \ x \ w \ y \ w \ r \ j \ y \ l \ \tilde{w} \ n \ \tilde{y} \ \tilde{y}\]
Konyagi

(Data from Merrill, ms.)

\[
\begin{array}{|c|cccccccccccccc|}
\hline
P & p & t & c & k & f & Ɤ & x & b & d & ġ & g & w & ŋ & l & m & n & ŋ & ŋ \\
\hline
P & f & ŋ & Ɤ & x & f & Ɤ & x & w & ŋ & w & r & j & ŋ & l & ŋ & n & ŋ & ŋ \\
B & f & s & Ɤ & x & f & Ɤ & x & w & ŋ & w & r & j & ŋ & l & ŋ & n & ŋ & ŋ \\
\hline
\end{array}
\]
Konyagi

(Data from Merrill, ms.)

P  p  t  c  k  f  j  x  b  d  j  g  w  γ  l  m  n  ŉ  ĕ

P  f  ř  ň  x  f  ň  x  w  γ  w  r  j  γ  l  ŕ  n  ť  ţ

B  f  ř  ň  x  f  ň  x  w  γ  w  r  j  γ  l  ŕ  n  ţ  ţ

K  f  ř  s  x  f  s  x  w  l  j  /  w  /  l  ŕ  n  ţ  /
Konyagi

(Data from Merrill, ms.)

P | p t c k f j x b d j g w y l m n ň ŋ
P | f ř j x f j x w y w r j y l ŕ n ŭ ŭ
B | f s j x f j x w y w r j y l ŕ n ŭ ŭ
K | ř r s x f s x w l j / w / l ŕ n ŭ /

P | p t c k f j x b d j g w y l m n ň ŋ
Konyagi

(Data from Merrill, ms.)

P  |  p t c k f ʃ x b d ŋ g w ŭ l m n ř ř
P  |  f ř ʃ x f ʃ x w ŭ w r j ŭ l ř ř ŕ ř
B  |  f s ʃ x f ʃ x w ŭ w r j ŭ l ř ŕ ŕ ŕ
K  |  f r s x f s x w l j / w / l ř ŕ ŕ /

P  |  p t c k f ʃ x b d ŋ g w ŭ l m n ř ř
P  |  p t c k p c k b d ŋ g b g l m n ř ř
# Konyagi

(Data from Merrill, ms.)

| P | p t c k f j x b d j g w γ l m n ŋ ķ |
| P | p t c k f j x w ẏ w r j γ l ŋ n ỹ ţ |
| B | f s j x f j x w ẏ w r j γ l ŋ n ỹ ţ |
| K | f r s x f s x w l j / w / l ŋ n ỹ / |

| P | p t c k f j x b d j g w γ l m n ŋ ķ |
| P | p t c k p c k b d j g b g l m n ŋ ķ |
| B | p t c k p c k b d j g b g l m n ŋ ķ |

Gašper Beguš  
Unnatural Trends in the Lexicon: Diachrony and Synchrony

Harvard University — begus@fas.harvard.edu
# Konyagi

(Data from Merrill, ms.)

<table>
<thead>
<tr>
<th>P</th>
<th>p t c k f f ∫ x b d j g w y l m n ŋ ŋ</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>f r ∫ x f f ∫ x w y w r j y l ŋ n ŋ ŋ ŋ</td>
</tr>
<tr>
<td>B</td>
<td>f s ∫ x f f ∫ x w y w r j y l ŋ n ŋ ŋ ŋ</td>
</tr>
<tr>
<td>K</td>
<td>f r s x f s x w l j / w / l ŋ n ŋ</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>P</th>
<th>p t c k f f ∫ x b d j g w y l m n ŋ ŋ</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>p t c k p c k b d j g b g l m n ŋ ŋ</td>
</tr>
<tr>
<td>B</td>
<td>p t c k p c k b d j g b g l m n ŋ ŋ</td>
</tr>
<tr>
<td>K</td>
<td>p t c k p c k p t c k p k l m n ŋ ŋ</td>
</tr>
</tbody>
</table>
Konyagi

(Data from Merrill, ms.)

| P | p t c k f j x b d j g w y l m n ŋ ŋ |
| P | f ř j x f j x w y w r j y l ſ n ſ ſ |
| B | f s j x f j x w y w r j y l ſ n ſ ſ |
| K | f r s x f s x w l j / w / l ſ n ſ / |

P | p t c k f j x b d j g w y l m n ŋ ŋ |
| P | p t c k p c k b d j g b g l m n ŋ ŋ |
| B | p t c k p c k b d j g b g l m n ŋ ŋ |
| K | p t c k p c k p t c k p k l m n ŋ ŋ |
South Italian Dialects

- Such combination of two sound changes is **not limited** to stops
South Italian Dialects

- Such combination of two sound changes is not limited to stops
- When other segments are targeted, we have strong evidence for a stage with *complementary distribution* as well
South Italian Dialects

Ducibella (1934, 450):

<table>
<thead>
<tr>
<th>S.-Ital. dial.</th>
<th>Standard</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>antšilu</td>
<td>andĝelo</td>
<td>‘angel’</td>
</tr>
<tr>
<td>pintširi</td>
<td>pindĝere</td>
<td>‘push’</td>
</tr>
<tr>
<td>chainširi</td>
<td>planĝere</td>
<td>‘to cry’</td>
</tr>
<tr>
<td>fintširi</td>
<td>findĝere</td>
<td>‘to feign’</td>
</tr>
<tr>
<td>tintširi</td>
<td>tindĝere</td>
<td>‘to dye’</td>
</tr>
</tbody>
</table>
South Italian Dialects

- Ducibella (1934, 450):

<table>
<thead>
<tr>
<th>S.-Ital. dial.</th>
<th>Standard</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ant⁵gilu</td>
<td>angelo</td>
<td>‘angel’</td>
</tr>
<tr>
<td>pin⁵giri</td>
<td>pindgere</td>
<td>‘push’</td>
</tr>
<tr>
<td>chaid⁵giri</td>
<td>plandgere</td>
<td>‘to cry’</td>
</tr>
<tr>
<td>fin⁵giri</td>
<td>fingere</td>
<td>‘to feign’</td>
</tr>
<tr>
<td>tint⁵giri</td>
<td>tingere</td>
<td>‘to dye’</td>
</tr>
</tbody>
</table>
South Italian Dialects

- Ducibella (1934, 450):

<table>
<thead>
<tr>
<th>S.-Ital. dial.</th>
<th>Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>anfìlu</td>
<td>andëlo</td>
</tr>
<tr>
<td>pintìri</td>
<td>pintère</td>
</tr>
<tr>
<td>chaintìri</td>
<td>planère</td>
</tr>
<tr>
<td>fintìri</td>
<td>fìnère</td>
</tr>
<tr>
<td>tintìri</td>
<td>tinère</td>
</tr>
</tbody>
</table>

- No attempts to explain the distribution
South Italian Dialects

- Crucially: elsewhere $\acute{\eta}$3 gets deoccluded to $\eta$3
South Italian Dialects

- Crucially: elsewhere $\dot{q}3$ gets deoccluded to $3$
- Complementary distribution: $faq3ina > fa3ina$
South Italian Dialects

- Crucially: elsewhere $d\_z$ gets deoccluded to $z$
- Complementary distribution: $fad\_zina > fa\_zina$

- Proposal:
  **Stage 1** Deocclusion of voiced affricate except after nasals
  *(complementary distribution)*
  \[DZ > D / [-\text{nas}]\]
South Italian Dialects

- Crucially: elsewhere ꢣ gets deoccluded to ꢤ
- Complementary distribution: \( \text{fad}_\text{ḏ} \text{zina} > \text{fa}_\text{ḏ} \text{zina} \)

Proposal:

**Stage 1**  Deocclusion of voiced affricate except after nasals

(complementary distribution)

\[ \text{DZ} > \text{D} / [\text{-nas}] \]

**Stage 2**  Unconditioned devoicing of voiced affricates

\[ \text{DZ} > \text{TS} \]
Buginese, Murik, Land Dayak

- Blust (2013):

<table>
<thead>
<tr>
<th>Proto-SS</th>
<th>Buginese</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>*bemba</td>
<td>bempa</td>
<td>‘water jar’</td>
</tr>
<tr>
<td>*lambuk</td>
<td>lampuʔ</td>
<td>‘pound rice’</td>
</tr>
<tr>
<td>*limboŋ</td>
<td>lempoŋ</td>
<td>‘deep water’</td>
</tr>
<tr>
<td>*rambu</td>
<td>rampu</td>
<td>‘fringe’</td>
</tr>
<tr>
<td>*rumbia</td>
<td>rumpia</td>
<td>‘sago palm’</td>
</tr>
<tr>
<td>*tambiŋ</td>
<td>tampiŋ</td>
<td>‘addition to a house’</td>
</tr>
<tr>
<td>*barumbun</td>
<td>warumpuŋ</td>
<td>‘a color pattern’</td>
</tr>
<tr>
<td>*bumbun</td>
<td>wumpuŋ</td>
<td>‘heap up’</td>
</tr>
<tr>
<td>*aŋgəp</td>
<td>aŋkəʔ</td>
<td>‘price’</td>
</tr>
<tr>
<td>*anjap</td>
<td>ancəʔ</td>
<td>‘offerings to spirits’</td>
</tr>
</tbody>
</table>
## Buginese, Murik, Land Dayak

- **Blust (2013):**

<table>
<thead>
<tr>
<th>Proto-SS</th>
<th>Buginese</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>*bemba</td>
<td>bempa</td>
<td>‘water jar’</td>
</tr>
<tr>
<td>*lambuk</td>
<td>lambu?</td>
<td>‘pound rice’</td>
</tr>
<tr>
<td>*limboŋ</td>
<td>lempoŋ</td>
<td>‘deep water’</td>
</tr>
<tr>
<td>*rambu</td>
<td>rampu</td>
<td>‘fringe’</td>
</tr>
<tr>
<td>*rumbia</td>
<td>rumpia</td>
<td>‘sago palm’</td>
</tr>
<tr>
<td>*tambiŋ</td>
<td>tampiŋ</td>
<td>‘addition to a house’</td>
</tr>
<tr>
<td>*barumbun</td>
<td>warumpuŋ</td>
<td>‘a color pattern’</td>
</tr>
<tr>
<td>*bumbun</td>
<td>wumpuŋ</td>
<td>‘heap up’</td>
</tr>
<tr>
<td>*aŋgẹp</td>
<td>aŋkẹ?</td>
<td>‘price’</td>
</tr>
<tr>
<td>*anjap</td>
<td>ancẹ?</td>
<td>‘offerings to spirits’</td>
</tr>
</tbody>
</table>
## Buginese, Murik, Land Dayak

<table>
<thead>
<tr>
<th>Proto-KM</th>
<th>Murik</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>*kelembit</td>
<td>kələmpit</td>
<td>‘shield’</td>
</tr>
<tr>
<td>*umbuŋ</td>
<td>umpuŋ</td>
<td>‘ridge of a roof’</td>
</tr>
<tr>
<td>*lindem</td>
<td>lintəm</td>
<td>‘dark’</td>
</tr>
<tr>
<td>*mandaŋ</td>
<td>mantaŋ</td>
<td>‘to fly’</td>
</tr>
<tr>
<td>*tundek</td>
<td>tuntuk</td>
<td>‘beak of a bird’</td>
</tr>
<tr>
<td>*lindiŋ</td>
<td>lintiŋ</td>
<td>‘wall of a house’</td>
</tr>
<tr>
<td>*anderŋ</td>
<td>antəŋ</td>
<td>‘deaf’</td>
</tr>
<tr>
<td>*pindaŋ</td>
<td>pintaŋ</td>
<td>‘blossom’</td>
</tr>
<tr>
<td>*pendan</td>
<td>pəntan</td>
<td>‘small fruit bat’</td>
</tr>
<tr>
<td>*ŋəgi</td>
<td>nʧi</td>
<td>‘one’</td>
</tr>
<tr>
<td>*mendŋat</td>
<td>məntʧat</td>
<td>‘pull’</td>
</tr>
</tbody>
</table>
## Buginese, Murik, Land Dayak

<table>
<thead>
<tr>
<th>Proto-KM</th>
<th>Murik</th>
<th>‘Meaning’</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>kelembit</em></td>
<td>kələmpit</td>
<td>‘shield’</td>
</tr>
<tr>
<td><em>umbuŋ</em></td>
<td>uŋpuŋ</td>
<td>‘ridge of a roof’</td>
</tr>
<tr>
<td><em>lindem</em></td>
<td>lintəm</td>
<td>‘dark’</td>
</tr>
<tr>
<td><em>mandaŋ</em></td>
<td>mantəŋ</td>
<td>‘to fly’</td>
</tr>
<tr>
<td><em>tundek</em></td>
<td>tuntuk</td>
<td>‘beak of a bird’</td>
</tr>
<tr>
<td><em>lindiŋ</em></td>
<td>lintiŋ</td>
<td>‘wall of a house’</td>
</tr>
<tr>
<td><em>andeŋ</em></td>
<td>antəŋ</td>
<td>‘deaf’</td>
</tr>
<tr>
<td><em>pindaŋ</em></td>
<td>pintəŋ</td>
<td>‘blossom’</td>
</tr>
<tr>
<td><em>pendan</em></td>
<td>pəntəŋ</td>
<td>‘small fruit bat’</td>
</tr>
<tr>
<td><em>nɔŋi</em></td>
<td>nɔŋi</td>
<td>‘one’</td>
</tr>
<tr>
<td><em>mendŋat</em></td>
<td>məntʃat</td>
<td>‘pull’</td>
</tr>
</tbody>
</table>

Gašper Beguš  
Harvard University — begus@fas.harvard.edu

Unnatural Trends in the Lexicon: Diachrony and Synchrony
Buginese, Murik, Land Dayak

- Here too clear traces of a stage with **complementary distribution**
- Bug. *b > w and *d > r
- Mur. *d > r initially and intervocally, but remains a stop post-nasally
Buginese, Murik, Land Dayak

- Here too clear traces of a stage with complementary distribution
- Bug. *b > w and *d > r
- Mur. *d > r initially and intervocalically, but remains a stop post-nasally

**Stage 1** Fricativization of voiced stops except after nasals (complementary distribution)

\[ D > \bar{D} / [-nas] \]

**Stage 2** Unconditioned devoicing of voiced stops

\[ D > T \]

**Stage 3** Velar voiced fricative undergoes occlusion, w and r stay unchanged

\[ ^*\gamma > g \]
PND as a synchronic process

- Coetzee and Pretorius (2010):
  - Complete neutralization
  - Applies with the same rate to nonce-words
- Also in Buginese (Sirk 1983:35-37)
All known cases of PND point to common patterns: a unified account of PND
Interim Summary

- All known cases of PND point to common patterns: a unified account of PND
- PND is a side effect of two or three phonetically well-motivated and well-attested sound changes:
  
  \[ D > Z \/ \ [-nas] \]
  
  \[ D > T \]
  
  \[ Z > D \]
- Sound change targets surface allophones
Interim Summary

- All known cases of PND point to common patterns: a unified account of PND

- PND is a side effect of two or three phonetically well-motivated and well-attested sound changes:

\[ D > Z / [-nas] \]

\[ D > T \]

\[ Z > D \]

- Sound change targets surface allophones

- Important consequences: it is no longer necessary to assume that sound change operates in phonetically unnatural direction. Likewise, we no longer need to assume PND is a phonetically natural process
Interim Summary

- **Blurring Process**
  - A historical model for explaining unnatural phenomena
  - Addition to current strategies (hypercorrection)
    - (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
Models

- Blurring process
- \( A > B / X = UPT \)
- How do we get \( B > A / X \)
  - BLURRING CYCLE
  - \( B > C / -X \)
  - \( B > A \)
  - \( C > B \)
Models

- Blurring process
- $A > B / X = UPT$
- How do we get $B > A / X$
  
  **BLURRING CYCLE**
  
  $B > C / -X$
  $B > A$
  $C > B$
  
  **BLURRING CHAIN**
  
  $B > C / X$
  $C > D$
  $D > A$
Outline

1. Introduction
2. Data
3. Background
4. Diachrony
5. Typology
6. Theoretical implications
7. Conclusion
Berawan
Berawan dialects

- In Berawan dialects labials and velars undergo **devoicing**, as reported in Blust (2013)
- Labials additionally change place of articulation (but only intervocalically!)
Berawan dialects

- g (from r) > k / V__V

  *bigiu > bikiw
  *gigiəq > giki?
  *magi > maki
  *igiəŋ > ikiŋ
  *ugat > ikit
Berawan dialects

- g (from r) $\rightarrow$ k / V__V
  - *bigiu $\rightarrow$ bikiw
  - *gigiəq $\rightarrow$ giki?
  - *magi $\rightarrow$ maki
  - *igiəŋ $\rightarrow$ ikịŋ
  - *ugat $\rightarrow$ ikịt
Berawan dialects

- **g (from r) > k / V__V**
  - *bigiu > bikiw*
  - *gigiəq > giki?
  - *magi > maki*
  - *giŋəŋ > ikŋ*
  - *ugat > ikit*

- **b > k / V__V**
  - *abiəŋ > akiŋ*
  - *bibì > biki*
  - *bəlibiəw > bəlikiəw*
  - *bibuy > bikuy*
  - *dibiəŋ > dikin*
Berawan dialects

- **g (from ɾ)** $\rightarrow$ **k / V__V**
  - *bigiu* $\rightarrow$ bikiw
  - *gigiəq* $\rightarrow$ giki?
  - *magi* $\rightarrow$ maki
  - *igiəŋ* $\rightarrow$ ikiŋ
  - *ugat* $\rightarrow$ ikit

- **b** $\rightarrow$ **k / V__V**
  - *abiəŋ* $\rightarrow$ akiŋ
  - *bibi* $\rightarrow$ biki
  - *bəlibiəw* $\rightarrow$ bəlikiəw
  - *bibuy* $\rightarrow$ bikuy
  - *dibiəŋ* $\rightarrow$ dikin
Long Terawan

- Word-initially, stops remain voiced

\[
\begin{array}{lcl}
*gəm & > & gəm \\
*gigun & > & gikuŋ \\
*gimot & > & gimok \\
*gitaq & > & gitaʔ \\
*bitok & > & bitok \\
*buliən & > & bulin \\
*busak & > & busek \\
*buttan & > & buten \\
\end{array}
\]
Long Terawan

Word-initially, stops remain voiced

*ɡəm > ɡəm
*ɡigung > ɡikuŋ
*ɡimot > ɡimok
*ɡitaq > ɡitaʔ

*bitok > bitok
*buliʔen > bulin
*busak > busek
*buttan > buten
Crucial problem

- \( *b > k / V_\_V \)

<table>
<thead>
<tr>
<th>Scenario 1</th>
<th>Scenario 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ( b &gt; p )</td>
<td>( b &gt; g )</td>
</tr>
<tr>
<td>2. ( p &gt; k )</td>
<td>( g &gt; k )</td>
</tr>
</tbody>
</table>

- **Problem:** if devoicing occurs before change of place, why no change of the original \( *p \)?
Crucial problem

- **b > k / V__V**

  **Scenario 1**  **Scenario 2**

  1. b > p  b > g
  2. p > k  g > k

- Problem: if devoicing occurs before change of place, why no change of the original *p?

  *hapuy > apuy, not **akuy
Crucial problem

- **b > k / V__V**

<table>
<thead>
<tr>
<th>Scenario 1</th>
<th>Scenario 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. b &gt; p</td>
<td>b &gt; g</td>
</tr>
<tr>
<td>2. p &gt; k</td>
<td>g &gt; k</td>
</tr>
</tbody>
</table>

- Problem: if devoicing occurs before change of place, why no change of the original *p?  
  *hapuy > apuy, not **akuy

- If change of place first, why only word-internally?
Berawan dialects (Blust 2005)

- Berawan stops

\[ D > T / V___V \]

<table>
<thead>
<tr>
<th></th>
<th>#___</th>
<th>V___V</th>
</tr>
</thead>
<tbody>
<tr>
<td>*b</td>
<td>b</td>
<td>k</td>
</tr>
<tr>
<td>*d</td>
<td>d</td>
<td>r</td>
</tr>
<tr>
<td>*g</td>
<td>g</td>
<td>k</td>
</tr>
</tbody>
</table>
Berawan dialects (Blust 2005)

- Berawan stops

\[ D > T / V\_\_\_V \]

<table>
<thead>
<tr>
<th></th>
<th>#__</th>
<th>V__V</th>
</tr>
</thead>
<tbody>
<tr>
<td>*b</td>
<td>b</td>
<td>k</td>
</tr>
<tr>
<td>*d</td>
<td>d</td>
<td>r</td>
</tr>
<tr>
<td>*g</td>
<td>g</td>
<td>k</td>
</tr>
</tbody>
</table>
Berawan dialects (Blust 2005)

- Berawan stops

\[
D > T / V \_\_V
\]

<table>
<thead>
<tr>
<th></th>
<th>#___</th>
<th>V___V</th>
</tr>
</thead>
<tbody>
<tr>
<td>*b</td>
<td>b</td>
<td>k</td>
</tr>
<tr>
<td>*d</td>
<td>d</td>
<td>r</td>
</tr>
<tr>
<td>*g</td>
<td>g</td>
<td>k</td>
</tr>
</tbody>
</table>
Berawan dialects (Blust 2005)

- Blurring chain

<table>
<thead>
<tr>
<th>Blurring Chain</th>
<th>Berawan</th>
</tr>
</thead>
<tbody>
<tr>
<td>B &gt; C / X</td>
<td>D &gt; Z / V__V</td>
</tr>
<tr>
<td>C &gt; D</td>
<td>Z &gt; S</td>
</tr>
<tr>
<td>D &gt; A</td>
<td>S &gt; T</td>
</tr>
</tbody>
</table>
Berawan dialects

- Berawan stops: stage 1

<table>
<thead>
<tr>
<th>Pre-Berawan</th>
<th>Berawan</th>
</tr>
</thead>
<tbody>
<tr>
<td>#_V_V</td>
<td>#_V_V</td>
</tr>
<tr>
<td>b</td>
<td>b β</td>
</tr>
<tr>
<td>d</td>
<td>d ḁ</td>
</tr>
<tr>
<td>g</td>
<td>g γ</td>
</tr>
</tbody>
</table>
Berawan dialects

- Berawan stops: stage 1a

<table>
<thead>
<tr>
<th>Pre-Berawan</th>
<th>Berawan</th>
</tr>
</thead>
<tbody>
<tr>
<td>#_</td>
<td>V_V</td>
</tr>
<tr>
<td>b</td>
<td>b β</td>
</tr>
<tr>
<td>d</td>
<td>d r</td>
</tr>
<tr>
<td>g</td>
<td>g γ</td>
</tr>
</tbody>
</table>
Berawan dialects

- Berawan stops: stage 2

<table>
<thead>
<tr>
<th>Pre-Berawan</th>
<th>Berawan</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>b φ</td>
</tr>
<tr>
<td>d</td>
<td>d r</td>
</tr>
<tr>
<td>g</td>
<td>g x</td>
</tr>
</tbody>
</table>
Berawan dialects

- Berawan stops: stage 3

<table>
<thead>
<tr>
<th>Pre-Berawan</th>
<th>Berawan</th>
</tr>
</thead>
<tbody>
<tr>
<td>#_ V_V</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>g</td>
<td>g</td>
</tr>
</tbody>
</table>

_— Gašper Beguš, Harvard University — begus@fas.harvard.edu_
Berawan dialects

- Berawan stops: stage 4

<table>
<thead>
<tr>
<th>Pre-Berawan</th>
<th>Berawan</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>g</td>
<td>g</td>
</tr>
</tbody>
</table>

#_ V_V
Phonetics

- *babuj > *bibuj > *biϕuj > *bixuj > [bikuj]
- Explains dissymmetry: alveolar — labial, velar
- Solves the chronology problem

**Scenario 1**
1. b > p
2. p > k

**Scenario 2**
1. b > g
2. g > k

- Better motivates change of place: [ϕ] > [x] easier to motivate than [b] > [g]
- Natural sound changes
Explanations

**BLURRING CHAIN**

| B > C / X | T > S / [−nas, −#][−cons] |
| C > D     | S > Z                   |
| D > A     | Z > D                   |

Tarma Quechua
Explanation

Tarma Native

variable

Voiced
Voiceless

N_  V_V  R_  T_

0%  25%  50%  75%  100%

Gašper Beguš
Harvard University — begus@fas.harvard.edu
Unnatural Trends in the Lexicon: Diachrony and Synchrony
Explanation

- T > S / [−nas,−#]→[−cons]
  - *aptay > [haxʷtay]
  - *upyay > [uxyay]
  - *paki > *pʰaki > Imbabura Quechua [faki]
  - Voiced stops still surface as fricatives

- S > Z / V __
  - Voicing of fricatives prevocally is a common change
  - cops amagats [bz]
  - <uxδa-> [uxða-] < *uxθa- < *uktʰa-

- Z > D
  - Not categorical: voiced stops still surface as fricatives
  - Variation in voiceless series: [ϕlawta], [plauta] (Puente Baldoceda 1977)
Phonetics

- Stops surface as fricatives (apparent free variation)
[atbi]
Phonetics

- This strongly suggests that there was a stage in the development with **voiced fricatives**
- Occlusion to stops not operated categorically
## Typology

- A combination of sound changes appears unconstrained
- A single sound change is constrained, combination unconstrained
- How to derive typology?
Minimal number of sound changes

- Note that unnatural processes always require three operating sound changes
- \( B > A / X \)
Minimal number of sound changes

- Note that unnatural processes always require three operating sound changes
  - B > A / X
  - B > C / X
  - C > A

At least three sound changes

While it is theoretically possible B > C and C > A, the last sound change would necessary involve two features to change, which in fact means two sound changes as C differs from B in one feature, which cannot be the feature producing unnaturalness, and A differs from B in one feature, which has to be the feature producing unnaturalness in a given context.

Crucial assumption: sound change is a change in one feature in a given environment.
Minimal number of sound changes

- Note that unnatural processes always require three operating sound changes
- \( B > A / X \)
- \( B > C / X \)
- \( C > A \)

- **At least three sound changes**
  - While it is theoretically possible \( B > C \) and \( C > A \), the last sound change would necessarily involve two features to change, which in fact means two sound changes as \( C \) differs from \( B \) in one feature, which cannot be the feature producing unnaturalness, and \( A \) differs from \( B \) in one feature, which has to be the feature producing unnaturalness in a given context

- Crucial assumption: sound change is a change in one feature in a given environment
Minimal number of sound changes

Minimal Sound Change Requirement (MSCR)

Minimally three sound changes have to operate in combination for an unnatural process to arise. Minimally two sound changes have to operate in combination for an unmotivated process to arise.
Typology

- Rare phenomena produced by rare sound changes
  (Blevins 2008)

- Morphological phenomena rare because of complex history
  (Harris 2003, 2008)
Typology

- Blurring process and MSCR directly predict typology
  - **Natural**
  - **Unmotivated**: Two sound changes, less frequent (telescoping, Wang 1968)
  - **Unnatural**: At least three sound changes, least frequent
- Low probability not the only reason
- As soon as an $B \rightarrow A / X$ arises, the inverse universal phonetic tendency $A > B / X$ will begin operating against it
- Confirmed by Tswana
Model

- The model can be quantified
- Poisson stochastic process
- We can calculate historical probabilities of processes (Fussell et al. 1976, Eid 2011)

\[
P(T_1) = \int_0^t f_1 t_1 dt_1 \times \int_{t_1}^t f_2 t_2 dt_2 \times \int_{t_2}^t f_3 t_3 dt_3 \times \ldots \times \int_{t_{n-1}}^t f_n t_n dt_n
\]

where

\[
f_i = \lambda_i e^{-\lambda_i t}
\]
Typology

- Easier to implement model

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

\[ P(Alt) = P(T_1 \cup T_2 \cup T_3 \cup \ldots \cup T_n) \]
Outline

1. Introduction
2. Data
3. Background
4. Diachrony
5. Typology
6. Theoretical implications
7. Conclusion
Theoretical implications

- **Harmony (HG)** → probability, but given richness of the base, we cannot derive a system in which the unnatural element is more frequent.

- **Faith** and a markedness constraint *X

- If equal weights, \( P(\text{UNNAT}) < .5 \)

- If either **Faith** or *X* have greater weights:
  a) If **Faith** > *X*: \( P(\text{UNNAT}) = .5 \)
  b) If **Faith** < *X*: \( P(\text{UNNAT}) < .5 \)

- Or:
  a) \( w(\text{IDENT-IO}(\text{voi})) > w(*\text{T#}): P([\text{T#}]) = P([\text{D#}]) = .5 \)
  b) \( w(*\text{T#}) > w(\text{IDENT-IO}(\text{voi})): P([\text{T#}]) > P([\text{D#}]) \)
Theoretical implications

- **Disagree, *VDV, *ND**
- Restricted Con undergenerates?
- Trends in the lexicon? Accidental gaps?
- How do we model unnatural trends in the lexicon?
- Experimental data
Future work

- Where to go from here?
Future work

- Where to go from here?
- More precise calculations of sound change probabilities_rates
- Channel and analytical bias not mutually exclusive:
- Grounds for experimental data
Conclusion

- Gradient phonotactics can be unnatural
- A model of development for unnatural processes:
  - Unnatural alternations
  - Unnatural phonotactics (gradience)
- A model for explaining typology within the channel bias
- Grounds for experimental research
References


Beguš, Gašper. 2015a. Post-nasal devoicing as a sound change. Presentation at the 89th Annual LSA Meeting, Portland, OR.

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


References


References


References


Thank you!