# A natural origin for unnatural gradient phonotactics

Gašper Beguš

Harvard University
*begus@fas.harvard.edu*

13th OCP, Budapest, Hungary
January 13, 2016



## Introduction

- Can gradient phonotactics be "unnatural"?
- What is unnatural?
- A model for explaining unnaturalness
- How to capture (unnatural) gradience theoretically?



## Outline





## Gradient phonotactics

- Two aspects of OT widely discussed: how to handle **naturalness** and **gradience**
- Little has been said about intersection of the two: **unnatural gradient phonotactics**
- Gradience implies naturalness?
- **Question:** Can gradience be unnatural?
- **Claim:** unnatural gradient phonotactics exists
- Tarma Quechua stop voicing



## Naturalness

- Both "naturalness" and "gradience" are to some degree confusing
- Traditionally, unnatural all processes that lack phonetic motivation
- This division misses one important aspect: phonetically **unmotivated** vs. going in the **opposite direction**



## Naturalness

- A new proposal:
  - **Natural**: phonetically grounded
  - **Unmotivated**: lack phonetic motivations
  - **Unnatural**: operating in the opposite direction from universal phonetic tendency



## Naturalness

- **Natural** processes phonetically motivated, typologically common, usually attested as minor phonetic tendencies cross-linguistically
- Unmotivated: lack motivation
- Most processes in the survey (Blevins 2008) unmotivated
  a) p → s/____i
  b) i → u/d____
- No universal phonetic tendency operating against these processes



## Naturalness

- **Unnatural**: operating against universal phonetic tendency (UPT)
- **UPT**: phonetically (articulatory) motivated, typologically common, attested as phonetic tendency in languages without phonological process. Reverse processes usually not attested
- Examples: post-nasal devoicing (attested), intervocalic devoicing, final voicing (unnatested, cf. Yu 2004)

## Phonotactics

- Most studies on **phonotactics** involve unmotivated restrictions (Hayes and White 2013, Albright 2009)
- Likewise, studies on **gradience** only include natural processes
- No cases of unnatural gradient phonotactics so far



## Gradient phonotactics

- Gradient phonotactics subject of in depth theoretical study only recently (Frisch et al. 2004, Antilla 2008, Coetzee and Pater 2008)
- Generally accepted that gradience needs to be encoded in grammar
- Shows necessity for weights in phonology



## Outline





## Data

- Tarma Quechua, a dialect of Quechua spoken in Tarma, Junín, Peru (Adelaar 1977, Nazarov 2008)
- **Stop voicing**
- Proto-Quechua, Pre-Tarma only voiceless stops
- Voicing occurs: intervocalically, post-consonantally, but not post-nasally
- Bilabial, velar series undergo voicing, dental remain voiceless



## Data

- From Adelaar 1977 and Nazarov 2008

$$p, k > b, g \;/\; C\_\_\_\_; C \neq N$$

$$p, k > b, g \;/\; V\_\_\_\_V$$

- Examples:

| | | |
|---|---|---|
| **#__** | *pirwa | pirwa |
| **N__** | *wampu- | wampu |
| **V__V** | *kupa- | kuba |
| | *kipu- | kipu |
| **Y,R,S,T__** | *takpa- | takba |

- In some words: lexical variation, not produtive



## Data

- Post-consonantally voicing after (Nazarov 2008): **t, ʧ, ʦ, k, s, ʃ, x, l, ʲ, r, j, w**

| **Pre-Tarma** | **Tarma** | **Pre-Tarma** | **Tarma** |
|---|---|---|---|
| aʧka | aʧga | maʦka | maʦga |
| aʦpa | aʦba | arku | argu |
| arku | argu | kuʧka | kuʧga |
| awkis | awgis | luʃpi | luʃbi |
| aypa | ayba | pilpa | pilba |
| ʧaspu | ʧasbu | luxpi | luxbi |
| ʧilpi | ʧilbi | mutki | mutgi |
| takpa | takba | tikpa | tikba |
| lutpi | lutbi | tikpi | tikbi |

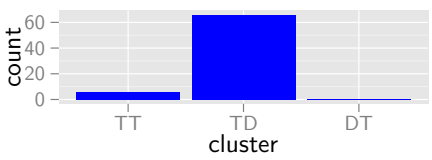


## Data

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

| **TT** | **TD** | **DT** | **DD** |
|---|---|---|---|
| 6 | 66 | 0 | 0 |

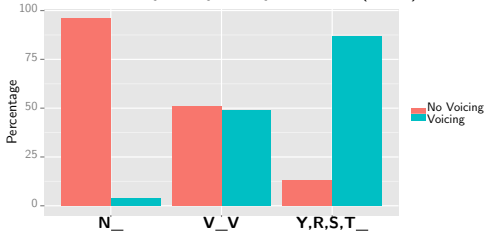
- Clusters that agree in voice preferred
- Pre-voicing preferred
- Tarma Quechua gradience in the **opposite direction**






## Data

- Even more surprising is the distribution
- Native vocabulary analyzed by Nazarov (2008)



## Data

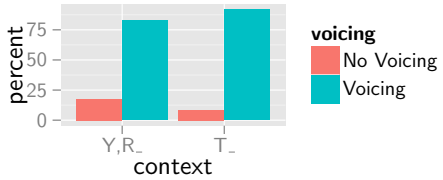
- Post-nasally voiced stops universally preferred
  (Hayes and Stivers 2000)
- After voiceless stops voiced stops universally dispreferred (voice disagree)
- Intervocalically, voiced stops universally preferred
- Gradience goes in the **opposite direction!**



## Data

- More voicing after T than after Y, R (Nazarov 2008)

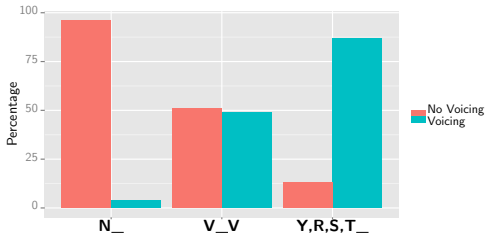
| | Y,R__ | T__ |
|---|---|---|
| Voiced | 66 | 66 |
| Voiceless | 14 | 6 |




- However, this is not significant $p < 0.10$



## Data




- $p < 0.00001$



## Explanations

- Not many explanations
- As with other cases of unnatural processes, the origins often sought in Ohala's **hypercorrection**
- Problems:
  - Why such distribution?
  - Why the asymmetry: dental vs. labial and velar?



## Outline





## Explaining unnatural processes

- Background: Post-nasal devoicing

  D > T / N__

- Unnatural process, attested as sound change in Yaghnobi and five other languages, as productive synchronic alternation in Tswana and Shekgalagari
- 3 sound changes, Yaghnobi preserves all three stages (Sogdian); Tswana clear dialectal evidence



## Blurring cycle

- Combination of sound changes: (Beguš 2015a)

  D > Z / [−*nas*]
  D > T
  Z > D



## Blurring cycle

- Universal phonetic tendency

  A > B / X

- Unnatural process

  B > A / X

- Blurring cycle

  B > C / Z; Z complement of X
  B > A
  C > B

  (a) a set of segments enters complementary distribution;
  (b) a sound change occurs that operates on the unchanged subset of those segments;
  (c) optionally, another sound change occurs that blurs the original complementary distribution.

## Blurring chain

- Another logically possible scenario

  B > A / X

- BLURRING CHAIN

  B > C / X
  C > D
  D > A



## Blurring cycle

- Note that unnatural processes always require three operating 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
- This requirement is the reason for why unnatural processes are less frequent than natural or unmotivated



## Berawan dialects (Blust 2005)

- Blurring chain explains **intervocalic devoicing** in Kiput and Berawan (Beguš 2015b)
- Unnatural sound change/phonotactics: voice contrastive initially
- Intervocalically neutralization in the dispreferred direction



## Berawan dialects (Blust 2005)

- Berawan stops

| | #__ | V__V |
|---|---|---|
| *b | b | k |
| *d | d | r |
| *g | g | k |

- ***b** > **k** / **V__V***

| | **Scenario 1** | **Scenario 2** |
|---|---|---|
| 1. | b > p | b > g |
| 2. | p > k | g > k |



## Berawan dialects (Blust 2005)

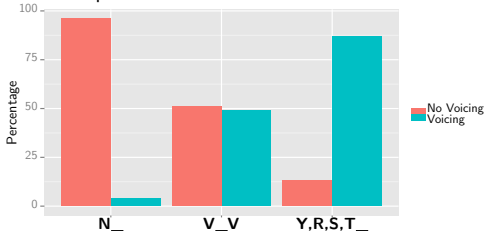
- Blurring chain

| BLURRING CHAIN | Berawan |
|---|---|
| B > C / X | D > Z / V__V |
| C > D | Z > S |
| D > A | S > T |



## Tarma Quechua

- How to explain such a distribution






## Tarma Quechua

- Claim: Tarma Quechua too underwent three sound changes



## Tarma Quechua

- If we accept blurring chain, the explanation becomes straight forward
- Fricativization: very common in consonant clusters, intervocalically
- Fricatives in post-nasal position highly dispreferred

## Tarma Quechua

- Fricativization in consonant clusters nearly categorical, intervocalically variation
- After that, voicing of fricatives occurs, which in pre-vocalic position is a common, motivated, and natural sound change
- Third sound change: fricative occlusion



## Tarma Quechua

| BLURRING CHAIN | Tarma Quechua |
|---|---|
| B > C / X | T > S / Y,R,S,T__, V__V |
| C > D | S > Z |
| D > A | Z > D |



## Advantages

- Asymmetry in place of articulation: velar and labial vs. dental
- Distribution of voicing:
  a) 4% post-nasally
  b) 49% intervocalically
  c) 87% in clusters

Three natural, well-motivated sound changes in combination gave rise to unnatural gradient phonotactics



## Outline



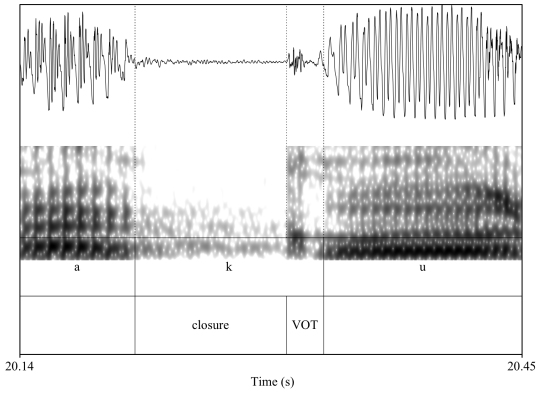


## Phonetics

- There exist strong phonetic evidence in favor of my proposal
- Old recordings by Willem Adelaar in Tarma, Peru
- Echoing disturbs the analysis
- The dialect highly endangered (Adelaar, p.c.)

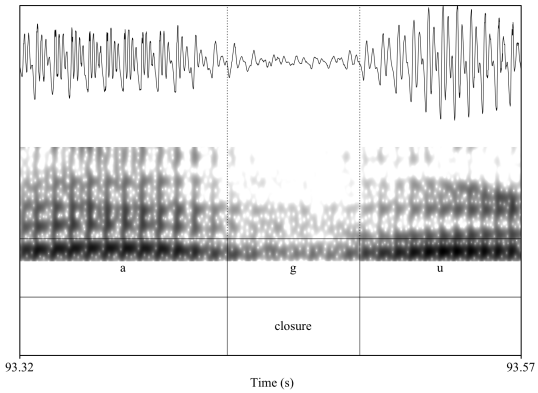


## [aku]



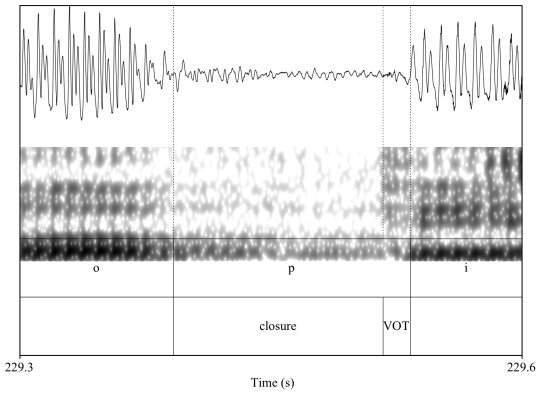



## [agu]

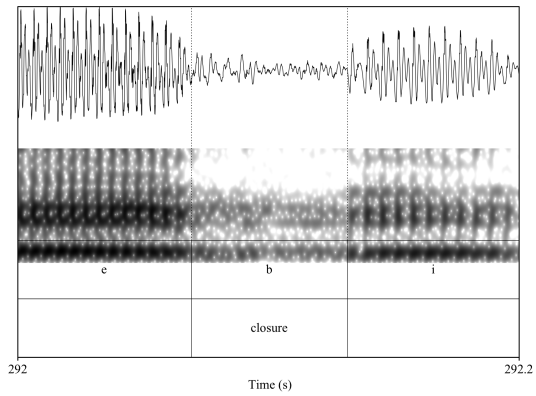





## [opi]

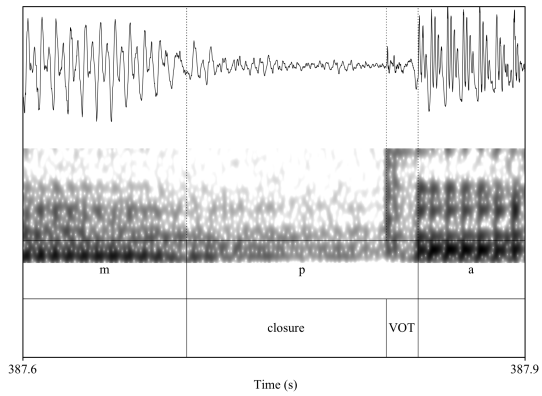


[ebi]



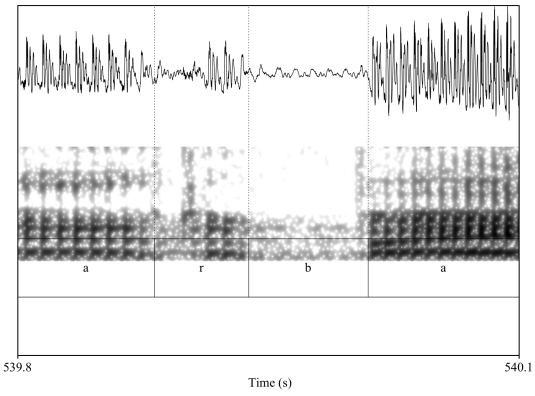

41 / 68

[mpa]



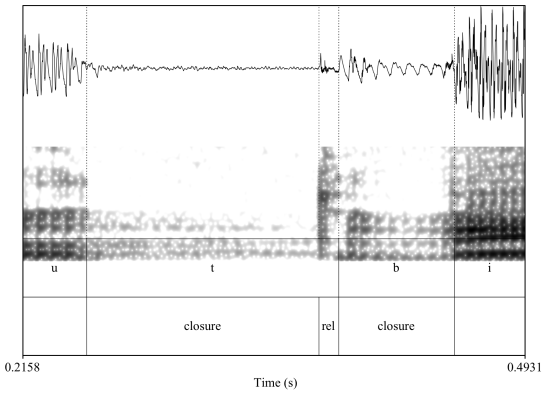

42 / 68

[arba]



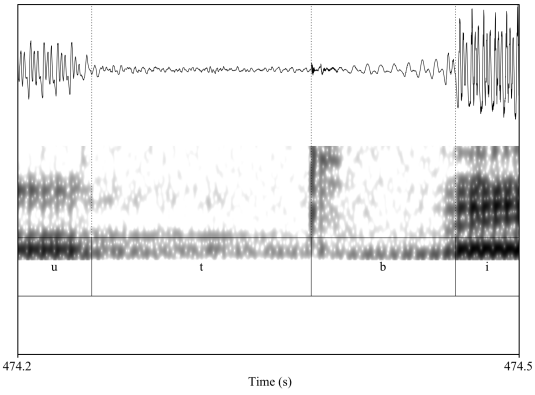

43 / 68

[utbi]



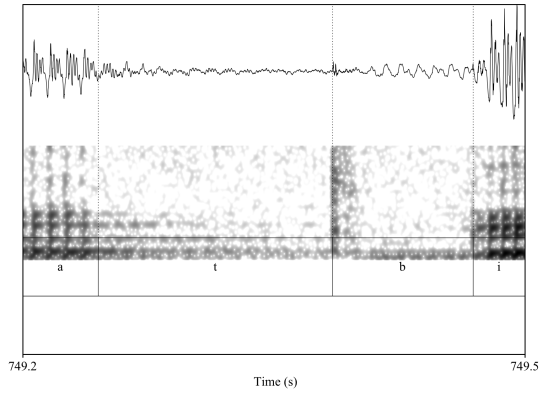

44 / 68

[utbi]



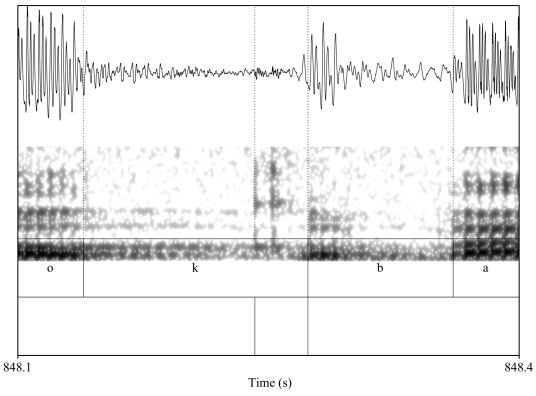

45 / 68

[atbi]



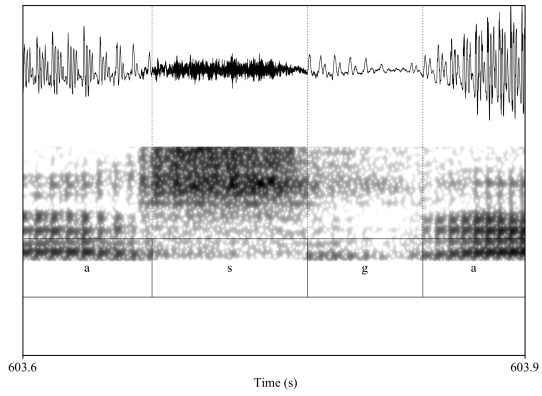

46 / 68

[okba]




47 / 68

[asga]



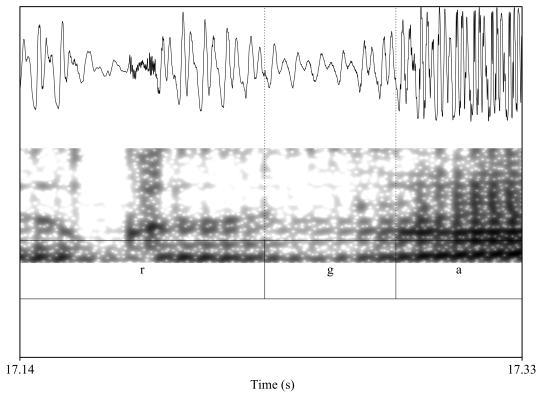

48 / 68

## Phonetics

- Stops surface as fricatives sometimes

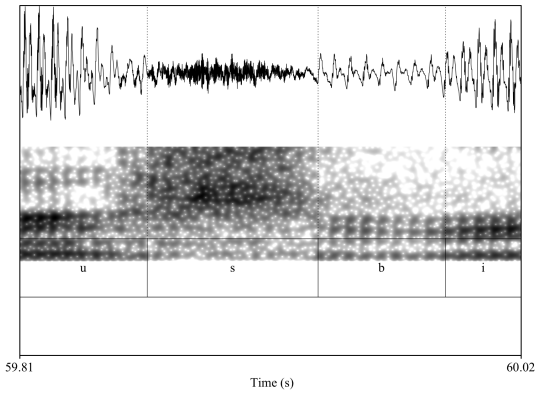


## [rga]



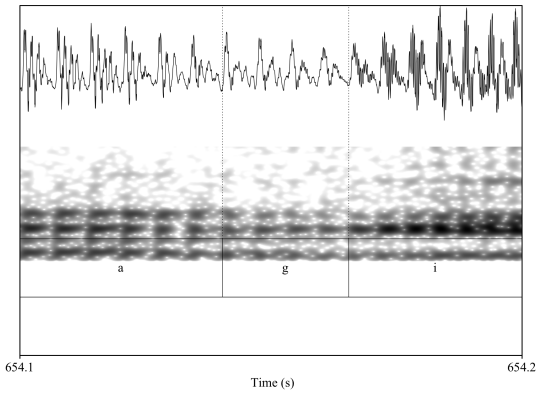



## [usbi]






## [agii]






## Phonetics

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



## Outline





## A problem for theory?

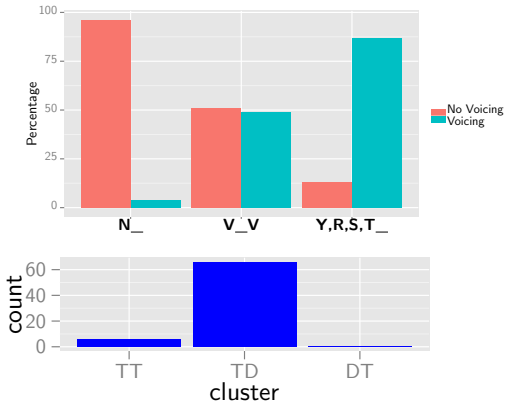
- It is generally agreed upon that gradient phonotactics has to be encoded in the grammar (Coetzee and Pater 2008, Antilla 2008)
- Various approaches for capturing gradience theoretically
- **The problem:** how to derive a system in which unnatural element is more common?



## A problem for theory?

- *Harmony* (HG) can be transformed to percentages, 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: $P(\text{UNNAT}) < .5$

## A problem for theory?






## A problem for theory?

- Should Con be restricted?
- *–X
- DisAgree?
- *VDV?
- **Problem:** how to encode these are rare?
- Not just Tarma Quechua



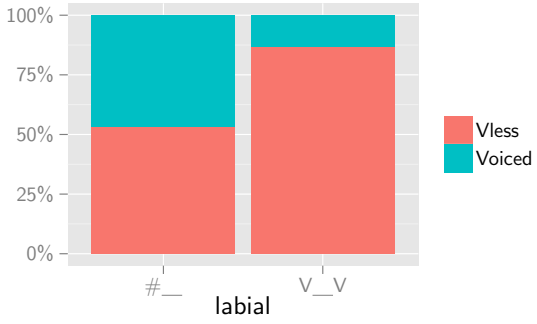
## A problem for theory?

- Berawan dialects (analysis based on data from Burkhardt 2014)

| | #__ | V__V | | #__ | V__V |
|---|---|---|---|---|---|
| [k] | 41 | 55 | [p] | 47 | 13 |
| [g] | 11 | 0 | [b] | 41 | 2 |
| **% voi** | 79 | 100 | | 53 | 86 |
| | $p < 0.001$ | | | $p < 0.05$ | |

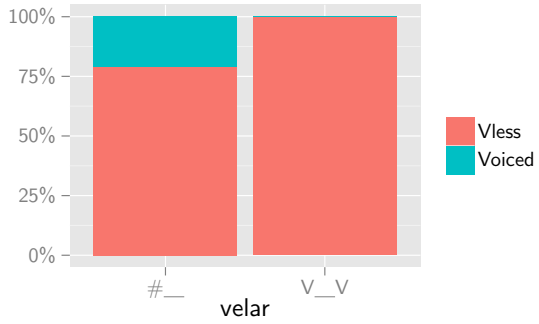


## Berawan






## Berawan






## A new proposal

- A new proposal: Inherently Weighted Constraints
- Both *X and *–X
- Constraints weighted, subject to normal distribution
- Derives such systems and encodes typology



## Outline





## Conclusion

- Gradient phonotactics can be unnatural
- Sound change restricted: combinations result in unnatural processes
- Blurring chain
- In non-negligible part of unnatural processes Blurring Cycle/Chain more successful than hypercorrection
- Theoretical implications: a new constraint architecture required

## References


Adelaar, Willem F. H. 1977. *Tarma Quechua: Grammar, texts, dictionary.* Lisse: The Peter de Ridder Press.

Adelaar, Willem F. H., and Muyske, Pieter C. 2004. *The Languages of the Andes*. Cambridge: CUP.

Albright, Adam. 2009. Feature-based generalisation as a source of gradient acceptability. *Phonology* 26(1): 9-41.

Anttila, Arto. 2008. Gradient phonotactics and the Complexity Hypothesis. *Natural Language & Linguistic Theory*, 26(4), 695-729.

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.

Blevins, Juliette. 2004. *Evolutionary Phonology*. Cambridge: CUP.

———. 2008. Natural and unnatural sound patterns : a pocket field guide. In Klaas Willems and Ludovic De Cuypere (eds.) *Naturalness and iconicity in language*. 121–148. Amsterdam: Benjamins.




## References


Blust, Robert. 2005. Must sound change be linguistically motivated? *Diachronica* 22 (2): 219–269.

Burkhardt, Jürgen M. The reconstruction of the phonology of Proto-Berawan. PhD Thesis, Johann-Wolfgang-Goethe-Universität zu Frankfurt am Main.

Coetzee, Andries. W., and Pater, Joe. 2008. Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic. *Natural Language & Linguistic Theory*, 26(2), 289-337.

Frisch, Stefan, Janet Pierrehumbert and Michael Broe. 2004. Similarity Avoidance and the OCP. *Natural Language and Linguistic Theory* 22. 179-228.

Hayes, Bruce, and Tanya Stivers. 2000. *Postnasal Voicing*. Unpublished ms.

Hayes, Bruce, and James White. 2013. Phonological naturalness and phonotactic learning. *Linguistic inquiry* 44(1): 45-75.

Hyman, Larry M. 2001. The Limits of Phonetic Determinism in Phonology: *NC Revisited. In *The Role of Speech Perception in Phonology*, ed. by Elizabeth Hume and Keith Johnson, 141–186. San Diego, CA: Academic Press.




## References


Nazarov, Aleksei. 2008. *Stop voicing in Tarma Quechua*. Ms., Leiden University.

Kaplan, Abby. 2010. *Phonology Shaped by Phonetics: The Case of Intervocalic Lenition*. PhD diss., University of California, Santa Cruz.

Kümmel, Martin. 2004. *Konsonantenwandel*. Wiesbaden: Reichert.

Ohala, John J. 1993. The phonetics of sound change. In *Historical Linguistics: Problems and Perspectives*, ed. by Charles Jones. 237–278. London: Longman.

Shiraishi, Hidetoshi. *Topics in Nivkh Phonology*. PhD Dissertation, Rijksuniversiteit Groningen.

Yu, A. C. L. 2004. Explaining final obstruent voicing in Lezgian: phonetics and history. *Language*, 80(1), 73-97.




*Thank you!*


* I would like to thank Kevin Ryan, Donca Steriade, and Edward Flemming for their useful comments and Willem Adelaar for his kind permission to analyze his recordings. Special thanks goes also to David Rockefeller Center for Latin American Studies at Harvard University for supporting my research. All mistakes are my own.