Post Nasal Devoicing as Opacity: A Problem for Natural Constraints

BRANDON PRICKETT

UNIVERSITY OF MASSACHUSETTS, AMHERST

35TH WEST COAST CONFERENCE ON FORMAL LINGUISTICS

Overview

1. Introduction
   i. Naturalness
   ii. Post-nasal devoicing
   iii. Duke of York opacity

2. Analysis
   i. Post-nasal devoicing as opacity
   ii. Learnability of opaque post-nasal devoicing
   iii. Word-final voicing as opacity
   iv. Learnability of opaque word-final voicing

3. Discussion

Why should constraints be natural?

- Limiting CON to a set of natural, universal constraints gives Optimality Theoretic approaches the ability to make strong typological predictions (Prince and Smolensky 2004 [1993], Hayes 1999).
  - Where natural means phonetically grounded, as in Hayes (1999).

- But are natural constraints **necessary** for correct typological predictions?
  - Recent computational approaches have modeled human language learning well without a requirement about a constraint's phonetic or typological naturalness (e.g. Hayes and Wilson 2008, Moreton et al. 2015).
  - Diachronic approaches have had success in explaining many typological trends (see, for example, Blevins 2004, Ohala 2005, Beguš submitted).
  - Weaker theories of naturalness (i.e. a naturalness bias) have also successfully predicted experimental results (see, for example, Wilson 2006, Hayes and White 2013).

- And are they **sufficient** for correctly predicting typology?
  - This question takes two forms:
  1. Do natural constraints underpredict typology?
  2. Do natural constraints overpredict typology?

- We're going to be looking at natural constraints' sufficiency in this presentation.
  - First we'll deal with underprediction, then we'll move on to overprediction.

Post-nasal devoicing

- Post-nasal devoicing (PND) is one pattern that has been proposed as evidence that natural constraints *underpredict* typology (Coetzee et al. 2007; see also Bach and Harms 1972 for more on "crazy" phonological processes).

  - PND in Tswana (from Coetzee et al. 2007)
    /m+**b**itsa/ → [m**p**itsa] '1st.Sg.Obj.call'
    /re+**b**itsa/ → [re**b**itsa] '1st.Pl.Obj.call'

- If we were to create a single constraint to motivate this process, the OT analysis would look something like this (see Hyman 2001):

  - *ND: Assign one * for every voiced obstruent that follows a nasal.

/mbitsa/	*ND	Ident(voice)
mbitsa	W*	L
→mpitsa		*

Naturalness of PND

- However, *ND is neither phonetically nor typologically natural.
  - The opposite process (post-nasal voicing) is much more common (see Pater 2004).
  - On the phonetics of PND: "nasal airflow leakage during stop articulation should promote…voicing" (Coetzee et al. 2007:861).
  - Although, see Coetzee et al. (2007) on how *ND could be motivated by perceptual factors.

- **So parallel OT with natural constraints underpredicts the presence of PND.**
  - This could be a problem with natural constraints, as Hyman (2001) suggests.
  - But it could also be a limitation of a strictly parallel version of OT.

- Can a non-parallel version of OT account for PND with only natural constraints?
  - Yes, Stratal OT (Booij 1996, Kiparsky 2000) can be used to represent PND using only natural constraints.
  - But we'll need to use Duke-of-York derivations (McCarthy 2003; Rubach 2003).

Duke-of-York opacity

- "Duke-of-York" derivations are a kind of phonological opacity in which segments that are changed in the process of a derivation return to their original form in the output (Pullum 1975).

  - "Oh, the grand old Duke of York,
    He had ten thousand men;
    He marched them up to the top of the hill,
    And he marched them down again." (English nursery rhyme)

  - In phonological terms:
    UR: /A/
    Rule 1:   A → B
    Rule 2:   B → A
    SR: [A]

- McCarthy (2003) talks about two kinds of Duke-of-York derivations:
  - Vacuous: nothing is dependent on the intermediate stage (like the above example).
  - Feeding: the intermediary stage feeds an independent process that would otherwise not be triggered.

Duke-of-York opacity (feeding)

- Feeding Duke-of-York derivation:

UR:	/AC/
Rule 1: A → B	BC
Rule 2: C → D/ B_	BD
Rule 3: B → A	AD
SR:	[AD]

- Real-life example of feeding Duke-of-York from Tiberian Hebrew (Prince 1975:87):

UR:	/bi+ktob/
Cluster break up: (∅ → V/ C_C)	bik**ə**tob
Spirantization: (T → S/V_V)	bi**x**ə**θ**ob
Schwa deletion: /ə/ → ∅/$\bar{V}C_a$_C_bV	bixθob
SR:	[bixθob]

- Rubach (2003) presents more evidence for feeding Duke-of-York derivations, citing polish velar palatalization and labial fusion as examples of processes that require such an analysis.
  - While McCarthy (2003) argues that, in general, Duke-of-York should be avoided, he does say that cases like Tiberian Hebrew (that act across morpheme boundaries) seem to require it.

PND as Duke-of-York Opacity

- Duke of York derivations are a synchronic version of the diachronic "telescoping" described by Wang (1968) and "blurring" proposed by Beguš (submitted).

- Dickens (1984) and Hyman (2001:163) use this diachronic opacity to explain how PND could have come about through a series of unrelated, natural diachronic changes.

Change	/mb/	/eb/
*D > Z/[-nasal]_	mb	eβ
*D > T	mp	eβ
*Z > D	mp	eb

- Beguš (submitted) shows how this process (and processes similar to it) can be independently motivated and used to explain essentially every case of PND.

Synchronic opacity and PND?

- If a feeding Duke-of-York derivation is used, post-nasal devoicing can be represented using only natural constraints.

- In the following analysis, I'll derive PND in a toy language that has no fricatives and no post-nasal voiced stops (this is for the sake of clarity; minor changes to the constraint set could make it applicable to a real-world example like Tswana).

- The natural constraints used in the derivation are described below:

*[+continuant]/N_	Assign one * for every continuant obstruent in the output that occurs after a nasal.
*[+voice,-continuant]	Assign one * for every voiced stop in the output.
*[+continuant]	Assign one * for every continuant obstruent.
Faith(F)	Assign one * for every segment in the input that has a different value for feature F in the output.

PND as opacity: /n+dad/ → [ntad]

- Stratum 1:
  - Avoidance of voiced stops repaired with frication, except post-nasally where it's repaired with devoicing.

/n+dad/	*[+voice,-cont.]	*[+cont.]/N_	Faith(voice)	Faith(cont.)	*[+cont.]
ndad	W**		L	L	L
ntad	W*		*	L	L
ndaz	W*		L	*	*
nzaz		W*	L	W**	W**
ntat			W**	L	L
→ntaz			*	*	*

- Stratum 2:
  - Avoidance of fricatives, repaired with fortition to stops.

ntaz	Faith(voice)	*[+cont.]	*[+cont.]/N_	*[+voice,-cont.]	Faith(cont.)
[ntaz]		W*		L	L
[nsaz]		W**	W*	L	*
→[ntad]				*	*

PND as opacity: /n+dad/ → [ntad]

- Stratum 1:
  - Avoidance of voiced stops repaired with frication, except post-nasally where it's repaired with devoicing.

ndad

ntaz

/n+dad/	*[+voice,-cont.]	*[+cont.]/N_	Faith(voice)	Faith(cont.)	*[+cont.]
ndad	W**		L	L	L
ntad	W*		*	L	L
ndaz	W*		L	*	*
nzaz		W*	L	W**	W**
ntat			W**	L	L
→ntaz			*	*	*

- Stratum 2:
  - Avoidance of fricatives, repaired with fortition to stops.

ntaz	Faith(voice)	*[+cont.]	*[+cont.]/N_	*[+voice,-cont.]	Faith(cont.)
[ntaz]		W*		L	L
[nsaz]		W**	W*	L	*
→[ntad]				*	*

PND as opacity: /n+dad/ → [ntad]

- Stratum 1:
  - Avoidance of voiced stops repaired with frication, except post-nasally where it's repaired with devoicing.

ndad

↓

ntaz

/n+dad/	*[+voice,-cont.]	*[+cont.]/N_	Faith(voice)	Faith(cont.)	*[+cont.]
ndad	W**		L	L	L
ntad	W*		*	L	L
ndaz	W*		L	*	*
nzaz		W*	L	W**	W**
ntat			W**	L	L
→ntaz			*	*	*

- Stratum 2:
  - Avoidance of fricatives, repaired with fortition to stops.

ntaz	Faith(voice)	*[+cont.]	*[+cont.]/N_	*[+voice,-cont.]	Faith(cont.)
[ntaz]		W*		L	L
[nsaz]		W**	W*	L	*
→[ntad]				*	*

PND as opacity: /n+dad/ → [ntad]

- Stratum 1:
  - Avoidance of voiced stops repaired with frication, except post-nasally where it's repaired with devoicing.

/n+dad/	*[+voice,-cont.]	*[+cont.]/N_	Faith(voice)	Faith(cont.)	*[+cont.]
ndad	W**		L	L	L
ntad	W*		*	L	L
ndaz	W*		L	*	*
nzaz		W*	L	W**	W**
ntat			W**	L	L
→ntaz			*	*	*

- Stratum 2:
  - Avoidance of fricatives, repaired with fortition to stops.

ntaz

↓

ntad

ntaz	Faith(voice)	*[+cont.]	*[+cont.]/N_	*[+voice,-cont.]	Faith(cont.)
[ntaz]		W*		L	L
[nsaz]		W**	W*	L	*
→[ntad]				*	*

PND as opacity: /n+dad/ → [ntad]

- Stratum 1:
  - Avoidance of voiced stops repaired with frication, except post-nasally where it's repaired with devoicing.

/n+dad/	*[+voice,-cont.]	*[+cont.]/N_	Faith(voice)	Faith(cont.)	*[+cont.]
ndad	W**		L	L	L
ntad	W*		*	L	L
ndaz	W*		L	*	*
nzaz		W*	L	W**	W**
ntat			W**	L	L
→ntaz			*	*	*

- Stratum 2:
  - Avoidance of fricatives, repaired with fortition to stops.

ntaz

↓

ntad

ntaz	Faith(voice)	*[+cont.]	*[+cont.]/N_	*[+voice,-cont.]	Faith(cont.)
[ntaz]		W*		L	L
[nsaz]		W**	W*	L	*
→[ntad]				*	*

Is opaque PND learnable?

- Representing PND as an opaque process demonstrates that an unnatural pattern can be predicted by a grammar with only natural constraints.

- However, this adds complexity to the overall representation, because one must use multiple strata to represent the pattern.
  - This could be seen as a hidden structure problem (Tesar 1998; see Nazarov 2016 on how Stratal OT learning can be viewed as a hidden structure problem).
  - Hidden structure patterns are not always learnable, since local optima can exist in the learning process (see, for example, Jarosz 2013).

- If learning of opaque PND is impossible, natural constraints would still be underpredicting the presence of PND.
  - (assuming that predicting an unlearnable pattern is equivalent to not predicting the pattern at all)

MaxEnt stratal learner

- I tested the learnability of opaque PND using the Maximum Entropy (MaxEnt) Stratal OT learner from Nazarov and Pater (to appear) .
  - See Goldwater and Johnson (2003) and Hayes and Wilson (2008) for more on MaxEnt learners.
  - See Staubs and Pater (2016) for more on learning MaxEnt grammars in a derivational framework.

- The learner maximizes the likelihood of the learning data my changing constraint weights (weights are analogous to OT constraint rankings).
  - Input tableaux give constraints violations for every input-output mapping.
  - Learning data is a table that gives probabilities for the different possible input-output mappings.
  - A regularization terms pressures constraints to have low weights.

- The learning algorithm minimizes KL-divergence between the probabilities predicted by the constraint weights and the input-output probabilities given to the learner.
  - The probability of a single input-output path is the product of each step in the derivation.
  - The probability of an input-output mapping is the sum of all the path probabilities that would end in that output.

MaxEnt Stratal Learning Example

UR	Stratum 1 Output	Stratum 2 Output (SR)
A	A	A
	B	B

- P(A→B) = ?

MaxEnt Stratal Learning Example

UR	Stratum 1 Output	Stratum 2 Output (SR)
A　———————→	A	A
	B	B

- $P(A \rightarrow B) = P(A \rightarrow A)_{\text{Stratum 1}} \cdots$

MaxEnt Stratal Learning Example

UR	Stratum 1 Output	Stratum 2 Output (SR)
A	A	A
	B	B

- $P(A \rightarrow B) = P(A \rightarrow A)_{Stratum\ 1} * P(A \rightarrow B)_{Stratum\ 2} \dots$

P(Path 1)

MaxEnt Stratal Learning Example

UR	Stratum 1 Output	Stratum 2 Output (SR)
A	A	A
	B	B

- $P(A \rightarrow B) = P(A \rightarrow A)_{\text{Stratum 1}} * P(A \rightarrow B)_{\text{Stratum 2}} + P(A \rightarrow B)_{\text{Stratum 1}} \ldots$

P(Path 1)

MaxEnt Stratal Learning Example

UR	Stratum 1 Output	Stratum 2 Output (SR)
A	A	A
	B	B

- $P(A \rightarrow B) = P(A \rightarrow A)_{Stratum\ 1} * P(A \rightarrow B)_{Stratum\ 2} + P(A \rightarrow B)_{Stratum\ 1} * P(B \rightarrow B)_{Stratum\ 2}$

P(Path 1)   P(Path 2)

MaxEnt Stratal Learning Example

UR	Stratum 1 Output	Stratum 2 Output (SR)
A	A	A
	B	B

- $P(A \rightarrow B) = P(A \rightarrow A)_{Stratum\ 1} * P(A \rightarrow B)_{Stratum\ 2} + P(A \rightarrow B)_{Stratum\ 1} * P(B \rightarrow B)_{Stratum\ 2}$

  $\underbrace{\qquad\qquad\qquad\qquad}_{P(Path\ 1)} \quad \underbrace{\qquad\qquad\qquad\qquad}_{P(Path\ 2)}$

- $P(A \rightarrow B) = P(Path\ 1) + P(Path\ 2)$

Learning PND

- Specifics:
  - The learner was run on data for an input-output mapping that demonstrated PND.
  - Learning data was categorical (i.e. /ndad/ → [ntad] had a probability of 1).
  - 100 simulations were run.
  - Different randomized initial constraint weights (uniform distribution from 0-10) were used on each of these runs.

- Learning data:

UR	SR	probability
ndad	ndad	0
ndad	nzad	0
ndad	nzaz	0
ndad	ntaz	0
ndad	ntat	0
ndad	ntad	1
ndad	ndaz	0

Learnability of opaque PND

- Results:
  - All 100 of the grammars (i.e. weighted constraint sets) produced by the learner assigned a probability of over 97% (M = 0.97453, SD = 0.00007) to the correct input-output mapping (/ndad/ → [ntad]).

- Example weights:

Stratum 1:

*[+voice,-cont.]	*[+cont.]/N_	Faith(voice)	Faith(cont.)	*[+cont.]
10.08317	10.08312	4.49227	0	0

- **Stratum 2:**

Faith(voice)	*[+cont.]	*[+cont.]/N_	*[+voice,-cont.]	Faith(cont.)
6.291652	5.607729	0	0	0

- The constraint weights above are analogous to the constraint rankings discussed earlier, with Stratum 2 crucially having a higher weight for ***[+continuant]** than for ***[+voice,-continuant]**.

Underprediction for PND? No.

- The above analysis showed that PND can be represented in a stratal framework, using only natural constraints and that this representation is learnable.
  - This is evidence that natural constraints don't underpredict typology in the case of PND, which has been one argument brought against them (see, for instance Hyman 2001).

- There's nothing in this theory to explain why PND is rare.
  - PND's rarity could be explained by factors that are not incorporated into the MaxEnt learner, like the multiple diachronic steps needed to produce PND (Beguš submitted).
  - Could also be the result of the specificity needed in the constraint weighting needed.

- Back to the main question: Is a theory of natural constraints sufficient?
  1. Do natural constraints underpredict typology?   No.
  2. **Do natural constraints overpredict typology?**

Word-final voicing

- Word final voicing has been argued to be a completely unattested pattern (Kiparsky 2006):
  - In a rule-based framework, it would look something like: T → D / _#
  - It's also phonetically unnatural, since voicing becomes more difficult to maintain the further one is in an utterance (see, for instance, Iverson and Salmons 2011).

- However, if we allow for Duke-of-York opacity (which was needed to represent PND), word-final voicing can be represented with only natural constraints.
  (see Blevins 2004 and Kiparsky 2006 for diachronic versions of this proposal)

- Possible diachronic origin for word-final voicing:
  - Original form   kat
  - Epenthesis   kat**a**
  - Intervocalic voicing   ka**d**a
  - Word-final vowel deletion kad
  - Final form   kad

Word-final voicing as opacity: /kat/ → [kad]

- Stratum 1:
  - Avoidance of codas, repaired with insertion. Avoidance of intervocalic voiceless stops, repaired with voicing.

/kat/	*Coda	*VTV	MAX	*V#	DEP	Faith(voice)
kat	W*			L	L	L
kad	W*			L	L	*
kata		W*		*	*	L
→kada				*	*	*
ka			W*	L		L

- Stratum 2:
  - Avoidance of word-final vowels, repaired with deletion.

kada	*V#	Faith(voice)	DEP	*VTV	MAX	*Coda
[kada]	W*				L	L
[kata]	W*	W*		W*	L	L
[kat]		W*			*	*
→[kad]					*	*
[ka]	W*				W**	L

Word-final voicing as opacity: /kat/ → [kad]

- Stratum 1:
  - Avoidance of codas, repaired with insertion. Avoidance of intervocalic voiceless stops, repaired with voicing.

kat

↓

kad**a**

/kat/	*Coda	*VTV	MAX	*V#	DEP	Faith(voice)
kat	W*			L	L	L
kad	W*			L	L	*
kata		W*		*	*	L
→kada				*	*	*
ka			W*	L		L

- Stratum 2:
  - Avoidance of word-final vowels, repaired with deletion.

kada	*V#	Faith(voice)	DEP	*VTV	MAX	*Coda
[kada]	W*				L	L
[kata]	W*	W*		W*	L	L
[kat]		W*			*	*
→[kad]					*	*
[ka]	W*				W**	L

Word-final voicing as opacity: /kat/ → [kad]

- Stratum 1:
  - Avoidance of codas, repaired with insertion. Avoidance of intervocalic voiceless stops, repaired with voicing.

kat

↓

kada

/kat/	*Coda	*VTV	MAX	*V#	DEP	Faith(voice)
kat	W*			L	L	L
kad	W*			L	L	*
kata		W*		*	*	L
→kada				*	*	*
ka			W*	L		L

- Stratum 2:
  - Avoidance of word-final vowels, repaired with deletion.

kada	*V#	Faith(voice)	DEP	*VTV	MAX	*Coda
[kada]	W*				L	L
[kata]	W*	W*		W*	L	L
[kat]		W*			*	*
→[kad]					*	*
[ka]	W*				W**	L

Word-final voicing as opacity: /kat/ → [kad]

- Stratum 1:
  - Avoidance of codas, repaired with insertion. Avoidance of intervocalic voiceless stops, repaired with voicing.

/kat/	*Coda	*VTV	MAX	*V#	DEP	Faith(voice)
kat	W*			L	L	L
kad	W*			L	L	*
kata		W*		*	*	L
→kada				*	*	*
ka			W*	L		L

- Stratum 2:
  - Avoidance of word-final vowels, repaired with deletion.

kad**a**

↓

kad

kada	*V#	Faith(voice)	DEP	*VTV	MAX	*Coda
[kada]	W*				L	L
[kata]	W*	W*		W*	L	L
[kat]		W*			*	*
→[kad]					*	*
[ka]	W*				W**	L

Word-final voicing as opacity: /kat/ → [kad]

- Stratum 1:
  - Avoidance of codas, repaired with insertion. Avoidance of intervocalic voiceless stops, repaired with voicing.

/kat/	*Coda	*VTV	MAX	*V#	DEP	Faith(voice)
kat	W*			L	L	L
kad	W*			L	L	*
kata		W*		*	*	L
→kada				*	*	*
ka			W*	L		L

- Stratum 2:
  - Avoidance of word-final vowels, repaired with deletion.

kada

↓

kad

kada	*V#	Faith(voice)	DEP	*VTV	MAX	*Coda
[kada]	W*				L	L
[kata]	W*	W*		W*	L	L
[kat]		W*			*	*
→[kad]					*	*
[ka]	W*				W**	L

Learnability of word-final voicing

- Results:
  - I ran the stratal MaxEnt learner 100 times with these constraints, the same kind of random initial weights, and training data that demonstrated categorical word-final voicing.
  - All 100 of the grammars (i.e. weighted constraint sets) produced by the learner assigned a probability of over 97% (M = 0.97918, SD = .0000009) to the correct input-output mapping (/kat/ → [kad]).
  - Example weights:
  Stratum 1:

*Coda	*VTV	MAX	*V#	DEP	Faith(voice)
5.628242	5.644507	5.090633	0	0	0

  - **Stratum 2:**

*V#	Faith(voice)	DEP	*VTV	MAX	*Coda
6.473508	6.348806	5.535137	0	0	0

  - The constraint weights above are analogous to the constraint rankings discussed earlier, with Stratum 2 crucially weighting ***V#** higher than **MAX**.

Discussion

- Is a theory of natural constraints sufficient?
    1. Do natural constraints underpredict typology?   <u>No.</u>   ☺
        - Natural constraints can represent PND in a learnable way.
    2. Do natural constraints overpredict typology?   <u>Yes.</u>   ☹
        - Natural constraints can represent word-final voicing in a learnable way.

- In order to represent PND, we introduced Duke-of-York derivations.

- However, Duke-of-York phonology proved to be too powerful: it also predicted the presence of word-final voicing.

- **This isn't a reason to completely abandon synchronic naturalness.**
    - However, it represents a problem with viewing a natural constraint set as the only tool for limiting typological predictions.

References

Becker, M., Ketrez, N., & Nevins, A. (2011). The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish laryngeal alternations. *Language, 87(1)*, 84-125.

Beguš, Gašper (submitted). *Post-nasal devoicing as a sound change*. MS: http://ling.auf.net/lingbuzz/003232

Blevins, J. (2004). *Evolutionary phonology: The emergence of sound patterns*. Cambridge University Press.

Coetzee, Andries, Susan Lin, & Rigardt Pretorius (2007). Post-nasal devoicing in Tswana. In *Proceedings of the 16th international congress of phonetic sciences*: 861-864.

Dickens, Patrick J (1984). The History of So-Called Strengthening in Tswana. *Journal of African Languages and Linguistics 6*:97–125.

Hale, M., & Reiss, C. (2000). "Substance Abuse" and" Dysfunctionalism": Current Trends in Phonology. *Linguistic Inquiry, 31(1)*, 157-169.

Hayes, B. P. (1999). Phonetically driven phonology. *Functionalism and formalism in linguistics, 1*, 243-285.

Jarosz, G. (2013). Learning with hidden structure in optimality theory and harmonic grammar: Beyond robust interpretive parsing. *Phonology, 30(01)*, 27-71.

Kiparsky, Paul (2000). "Opacity and Cyclicity." *The Linguistic Review 17 (2-4)*: 351–66.

Kiparsky, Paul (2006). Amphichronic linguistics vs. Evolutionary Phonology. *Theoretical Linguistics, 32*: 217–236.

McCarthy, J. J. (1993). A case of surface constraint violation. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, *38*(02), 169-195.

McCarthy, J. J. (2003). Sympathy, cumulativity, and the Duke-of-York gambit. *The syllable in optimality theory*, 23-76.

Mielke, Jeff (2008) *The Emergence of Distinctive Features*. Oxford: Oxford University Press.

Nazarov, Aleksei I. (2016). *Extending Hidden Structure Learning: Features, Opacity, and Exceptions*. Doctoral Dissertation. University Massachusetts, Amherst: http://scholarworks.umass.edu/dissertations_2/782

Nazarov, Aleksei and Joe Pater (to appear). Learning opacity in Stratal Maximum Entropy Grammar. *Phonology*.

Ohala, John J. (2005). *Phonetic Explanations for Sound Patterns: Implications for Grammars of Competence*. UC Berkeley: Department of Linguistics. Retrieved from: http://escholarship.org/uc/item/5m77t155

Parnell, M., & Amerman, J. D. (1977). Subjective evaluation of articulatory effort. *Journal of Speech, Language, and Hearing Research, 20(4)*, 644-652.

Pater, Joe (2004). Austronesian nasal substitution and other NC effects. In John J. McCarthy (ed.), *Optimality Theory in Phonology: A Reader*: 271-89. Malden, MA, and Oxford, UK: Blackwell.

Prince, Alan, and Paul Smolensky (2004 [1993]). *Optimality theory: Constraint interaction in generative grammar*. Oxford: Blackwell.

Rubach, J. (2003). Duke-of-York derivations in Polish. *Linguistic inquiry, 34(4)*, 601-629.

Tesar, Bruce. "An iterative strategy for language learning." *Lingua 104.1-2 (1998)*: 131-145.

Acknowledgments

Thanks to the students and professors in the class for which this analysis was developed, to the WCCFL reviewers for their comments, to the attendees of the joint UMass PRG-SSRG meeting for their feedback, to the members of the UMass 'Grant Group' for helping me to finalize the presentation, to Joe Pater for encouraging the project, Gaja Jarosz for helping me prepare it for a conference, to Aleksei Nazarov for lending me his MaxEnt learner and helpfully explaining how to use and describe it, and especially to Gašper Beguš (both for giving me feedback on an early version of this analysis and for inspiring it with his paper on diachronic blurring). The many errors remaining are all my own.

Constraint naturalness justifications

Constraint	Justification
*[+continuant]/N_	This constraint could be motivated by the difficulty in producing fricatives after nasals (Vaux 1998) and the fact that prenasalization is not as common for fricatives as it is for stops (Steriade 1993). Zulu, Greek, and Basque all show rules that could be motivated by this constraint (Mielke 2008).
*[+voice,-continuant]	"The default, normal state of obstruents is voiceless…" (Hayes 1994). 149 languages in Pbase (Mielke 2008) lack voiced stops. Over 10 of these have inventories with voiced fricatives (e.g. Assiniboine).
*[+continuant]	This could be motivated by the fact that fricatives require more effort than stops (Parnell and Amerman 1977). The language Agarabi has no fricatives, while no language in PBase (Mielke 2008) is listed as lacking stops.

Constraint naturalness justifications

Constraint	Justification
*V#	See McCarthy's (1993) Final-C constraint. Many languages avoid vowels word-finally. Examples from P-Base (Mielke 2008): Latvian, Kihungan, and Aymara. Examples from McCarthy (1993): Arabic, Yapese, and some dialects of English.
*VTV	"…forms that obey this constraint need not execute the laryngeal gestures needed to turn off voicing in a circumvoiced environment." (Hayes 1994) Examples from P-Base (Mielke 2008): Kwamera, Kalenjin, and Ao.
*Coda	See Prince and Smolensky's (2004 [1993]) –Cod constraint.

Word-final voicing using syllabification

Stratum 1

/ka.tat/	IdentOns(voice)	*Coda	MAX	*VTV	*V#	Ident(voice)	DEP
ka.tat		W*		*	L	L	L
ka.ta			W*	*	*	L	L
ka.da	W*		W*	L	*	*	L
ka.tata				W**	*	L	*
→ka.tada				*	*	*	*
ka.tad		W*		*	L	*	L
ka.data	W*			*	*	*	*
ka.dada	W*			L	*	W**	*
ka.dad	W*	W*		L	L	W**	L

Presence of a coda repaired with epenthesis, intervocalic [-voice] non-onsets repaired with voicing.

Word-final voicing using syllabification

Stratum 2

ka.tada	Ident(voice)	*V#	*Coda	*VTV	MAX	DEP	IdentOns(voice)
ka.tat	W*		*	*	*		
ka.ta		W*	L	*	**		
ka.da	W*	W*	L	L	**		
ka.tata	W*	W*	L	W**	L		
ka.tada		W*	L	*	L		
→ka.tad			*	*	*		
ka.data	W**	W*	L	*	L		W*
ka.dada	W*	W*	L	L	L		W*
ka.dad	W*		*	L	*		W*

Presence of a word-final vowel repaired with deletion.

Word-final voicing using syllabification

IdentOns(voice)	*Coda	MAX	*VTV	*V#	Ident(voice)	DEP
10.39135	5.509519	5.124203	4.840154	0	0	0

Ident(voice)	*V#	*Coda	*VTV	MAX	DEP	IdentOns(voice)
6.482325	6.430319782	0	0	0	0	0

After running the learner on this version of word-final voicing, I got the above weights, which are analogous to the necessary constraint rankings.