Post-nasal Devoicing as Opacity: A Problem for Natural Constraints

Brandon Prickett

1. Introduction

In constraint-based theories of phonology, phonetic and typological naturalness often motivate which constraints are proposed[1], with the assumption that natural constraints will lead to correct typological predictions (Prince and Smolensky 2004 [1993]).

Whether a naturalness criterion for constraints is necessary for limiting typology has been questioned. For example, recent computational work has mimicked human phonological learning with minimal restriction on which constraints are used (Hayes and Wilson 2008, Moreton et al. 2015), and weaker theories of naturalness (such as a naturalness bias on learning) have successfully predicted experimental results (Wilson 2006, Hayes and White 2013). Diachronic explanations of phonology have also been used to successfully limit typological predictions, without requiring hard limits on which phonological patterns are learnable (see, for example, Blevins 2004, Ohala 2005, Beguš, submitted).

In the current study, rather than exploring whether a naturalness restriction is *necessary* for restricting typology, I examine whether naturalness is *sufficient* for the task. Specifically, I ask whether a theory restricted to natural constraints underpredicts attested patterns, such as Post-Nasal Devoicing and if it overpredicts unattested patterns, such as Word-Final Voicing.

2. Background

2.1. Post-nasal Devoicing

Post-Nasal Devoicing (PND) is a phonological pattern in which underlyingly voiced plosives are devoiced after nasals. It is attested in a small number of languages, including Tswana (Coetzee et al. 2007; see Beguš, submitted for a more comprehensive list of languages that devoice post-nasally). An example of PND from Tswana is shown in (1):

(1) 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'

*Brandon Prickett, University of Massachusetts, Amherst, blprickett@outlook.com. Thanks to the students and professors in the class for which this analysis was developed, to the WCCFL reviewers and audience members 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 and Gaja Jarosz for advice on the project, 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).

[1] That is, constraints often must have phonetic motivation and must be shown to motivate typologically common patterns.

 Proceedings of the 35th West Coast Conference on Formal Linguistics, ed. Wm. G. Bennett et al., 331-340. Somerville, MA: Cascadilla Proceedings Project.

If one were to propose a single constraint to motivate PND, it would look something like (2):

(2) Simple OT analysis of PND (see Hyman 2001)

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

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

However, a constraint like *ND would be unnatural, since "nasal airflow leakage during stop articulation should promote…voicing" (Coetzee et al. 2007:861) and since the opposite of PND, Post-Nasal Voicing, is more typologically common (Pater 2004). Because of this, in strictly parallel OT, PND does present a counterargument to the idea that natural constraints are sufficient for predicting typology (Hyman 2001; see also Bach and Harms 1972 for more on "crazy" phonological processes). However, in the remainder of this paper I will show that Stratal Optimality Theory (Stratal OT; Booij 1996, Kiparsky 2000) *is* able to represent PND with only natural constraints, using "Duke-of-York" opacity.

2.2. Duke-of-York Opacity

Stratal OT (Booij 1996, Kiparsky 2000) is a version of optimality theory in which multiple strata each contribute to a phonological form's mapping from underlying representation (UR) to surface representation (SR). Each of these strata has its own constraints, input, and possible outputs, and corresponds to a particular step in a morpho-phonological derivation (e.g. lexical and post-lexical). The first stratum takes the UR as its input, and each subsequent stratum takes as input the output of the stratum before it. The final stratum's output is the SR of the form. Crucially, the ranking of constraints in each stratum is completely independent of the rankings of other strata.

Pullum (1975) first noted that when independent steps occur in a phonological derivation, a phenomenon known as "Duke-of-York" can occur (see McCarthy 2003 for more on this in constraint-based frameworks). This is a kind of phonological opacity in which segments that are changed in the process of a derivation return to their original form in the SR, sometimes triggering or obscuring other changes in the process. An abstract example of this is given in (3):

(3) Vacuous Duke-of-York

UR:		/A/
Step 1:	A → B	B
Step 2:	B → A	A
SR:		[A]

In the example above, Step 1 (which can be thought of as the first stratum in a Stratal OT derivation) changes the underlying form /A/ into B. B then becomes the input for the second Step (or stratum), which changes it back to A. This results in the surface representation being the same as the underlying representation, despite changes happening to the form over the course of the two intervening steps.

McCarthy (2003) discusses two kinds of Duke-of-York Opacity: "vacuous" and "feeding". Taken by itself, (3) is an example of a vacuous derivation, since /A/ appears to have not changed when just comparing the UR and SR, and since no other processes are involved. An example of a "feeding" Duke-of-York process is given below:

(4) Feeding Duke-of-York

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

In (4), not only is the mapping /A/ → B → [A] occurring (as in the vacuous process), but a separate process is also being fed in which /C/ maps to [D]. Since /C/ only undergoes this change in the context of B, and since it is not adjacent to a [B] in the SR, a feeding, Duke-of-York derivation is present. An example of this can be seen in Tiberian Hebrew when certain morphemes trigger vowel deletion, as seen in (5):

(5) Duke-of-York in Tiberian Hebrew (from Prince 1975:87)

UR:		/bi+ktob/
Cluster break-up:	(Ø → V/ C_C)	bik**ə**tob
Spirantization:	(T → S/V_V)	bi**xəθ**ob
Schwa deletion:	/ə/ → Ø/VC_a_C_bV	bixθob
SR:		[bixθob]

In the example above, a schwa is inserted to break up the cluster /kt/ and later deleted. By itself, this would constitute a vacuous Duke-of-York process. However, while the schwa is present it triggers spirantization so that /k/ and /t/ become [x] and [θ], respectively. In the SR, it appears that spirantization has occurred without the necessary environment (V_V), because the triggering schwa disappeared.

While McCarthy (2003) argued against Duke-of-York derivations generally, he suggested that they would be necessary for processes like (5) that occur across morpheme boundaries. Rubach (2003) also argued that Duke-of-York derivations are necessary, showing that two separate processes in Polish require them. In the remainder of this paper, I will assume that such derivations are possible for the phonological grammar and I will explore what this means for the predictions made by a theory restricted to natural constraints.

3. Analysis

3.1. Post-Nasal Devoicing as opacity

While it's true that a single ranking of natural constraints can't represent PND (see §2.1 for more on this), a Duke-of-York derivation in Stratal OT *can*. This can be achieved by breaking up PND into multiple, natural processes in a way that has long been used by historical linguists as a diachronic explanation for PND (see Dickens 1984, Hyman 2001:163). This is demonstrated below:

(6) Historical explanation for PND

	*mb	*eb
1) *D > Z/[-nasal]_	mb	eβ
2) *D > T	mp	eβ
3) *Z > D	mp	eb

In (6), the first diachronic change to occur in the proto-language is lenition of all stops that don't occur after nasals. The second change devoices all of the remaining stops, which only occur post-nasally due to the first change. Finally, the third change makes fricatives become stops,[2] obscuring the two changes that came before (creating an apparent case of PND). Beguš (submitted) showed how this process—which he called "blurring"[3]—can be independently motivated and used to explain the diachronic origins of every known case of PND.

This diachronic "blurring" is analogous to a synchronic Duke-of-York derivation, and (7) shows how the three historical changes above can be translated into a Stratal OT representation of PND in a toy language similar to Tswana.[4] See the Appendix for constraint descriptions and naturalness justifications.

---

[2] There are fricatives in Tswana's segment inventory, however they do not overlap in place of articulation with the stops that undergoer PND. Because of this, these fricatives do not affect my synchronic analysis.

[3] See also the diachronic "telescoping" described by Wang (1968).

[4] This is for the sake of clarity; minor changes to the constraint set could make it applicable to a real-world example like Tswana.

(7) PND as a Duke-of-York derivation

Stratum 1:

/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:

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

Stratum 1 mimics the first two historical changes. The constraint *[+voice,-cont.], which penalizes voiced stops, is satisfied with two different repair strategies: frication and devoicing. Frication of stops occurs everywhere except post-nasally, because of the constraint *[+cont.]/N_ that penalizes post-nasal fricatives. The result is that no voiced stops remain and that stops that have been devoiced occur only post-nasally. Stratum 2 obscures these changes in a way that's similar to the third diachronic change in (6). By having the constraint *[+cont.] ranked above *[+voice,-cont.] and Faith(cont.), this stratum causes all input fricatives to turn into stops. This creates a distribution of stops in the language that is consistent with PND, since only voiceless stops occur post-nasally, with voiced and voiceless stops occurring elsewhere.

The derivation above is similar to a feeding Duke-of-York derivation because the stop that doesn't occur post-nasally goes through a mapping of /d/ → z → [d], with the /d/ returning to its original form after becoming a fricative. Meanwhile, the mapping of /d/ → [t] occurs post-nasally, however the motivation for this change is hidden by the fact that voiced stops are reintroduced elsewhere in Stratum 2. This allows an apparent PND pattern to exist in the surface representation of the language, without ever positing an unnatural constraint like *ND. This demonstrates that a theory restricted to natural constraints can be made to predict PND (i.e. naturalness is not *underpredicting* in this instance).

3.2. Testing the learnability of opaque Post-Nasal Devoicing

The tableaux given in §3.1 demonstrate that Stratal OT can represent PND with only natural constraints. However, this adds complexity to the overall representation, since it requires multiple strata to represent the pattern. Nazarov (2016) showed how learning in stratal frameworks can be seen as a hidden structure problem, and hidden structure can cause issues for learnability (Tesar 1998; Jarosz 2013). Local optima can sometimes mean that "the learner is not guaranteed to converge" (Narzarov and Pater, to appear), meaning that patterns can sometimes be unlearnable. If this was the case for the stratal representation of PND given in §3.1, then natural constraints would still have a problem with underprediction, since predicting a pattern to be representable-but-also-unlearnable is effectively the same as not predicting it at all.

To test whether PND can be learned when it's represented using only natural constraints in a Stratal framework, I used the Maximum Entropy (MaxEnt) Stratal OT Learner from Nazarov and Pater (to appear). MaxEnt learners use constraint weights that are analogous to constraint rankings in classic OT, and the set the weights to values that will maximize the likelihood of the learning data they're given. See the Appendix or Nazarov and Pater (to appear) for a more detailed description of how the Stratal OT learner works.

I ran the MaxEnt learner on a set of learning data that demonstrated the UR-SR mapping /ndad/ → [ntad]. The learner was limited to the set of natural constraints used in the tableaux from (7) and was told to use two strata in its derivation. I ran it 100 times, with random initial constraint weights that were sampled each run from a uniform distribution of 0-10.

This resulted in the learner successfully acquiring opaque PND for all 100 runs. The UR-SR mapping /ndad/ → [ntad] was given an average probability of over 97% (M = 0.97453, SD = 0.00007). The constraint weights for one of these successfully learned grammars is given in (8). These weightings are analogous to the rankings in (7).

(8) Example constraint weights for PND

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

In Stratum 1, *[+voice,-cont.] has the highest weight, meaning that inputs to this stratum are very likely to change their voiced stops to something else. Because of the weight given to *[+cont.]/N_ and Faith(voice), the most likely changes are /d/ → t post-nasally and /d/ → z elsewhere. In Stratum 2, the crucial difference is that *[+cont.] has a higher weight than Faith(cont.). This causes the mapping of z → [d] to be very likely—the last change needed for the opaque PND process.

The results above show that PND as an opaque series of changes between the UR and SR, motivated by natural constraints, is consistently learnable. This is evidence that natural constraints don't underpredict the typology in the case of PND, which has been one argument brought against theories that have a naturalness criterion (see, for instance, Hyman 2001).

3.3. Word-final Voicing as opacity

The second question to address is whether natural constraints overpredict typology. That is, whether limiting a constraint set based on naturalness will keep a theory from predicting unattested phonological patterns. To do this, I used Word-Final Voicing (WFV) as an example of a completely unattested pattern (Kiparsky 2006). WFV is a process in which word-final obstruents surface as [+voice], regardless of their underlying value for the voicing feature, and is considered unnatural since voicing at the end of a word is more difficult to produce and perceive (see, for instance, Iverson and Salmons 2011).

Like PND in §3.1, WFV can be represented using natural constraints and Stratal OT. Again, I borrowed this line of logic from discussions about possible diachronic origins for WFV (Blevins 2004, Kiparsky 2006). This is demonstrated in (9):

(9) Hypothetical historical explanation for WFV

	*kat
1) *∅ > V/C_#	kata
2) *T > D/V_V	kada
3) *V > ∅/_#	kad

In the above example, the word *kat undergoes a series of three hypothetical diachronic changes to become the word [kad]. On the surface, this looks like WFV, however it's actually a series of other, more natural changes. The first change inserts a vowel after word-final consonants, the second change voices consonants intervocalically, and the last change deletes word-final vowels[5]. I analogize this diachronic proposal to a synchronic Duke-of-York derivation below (see the Appendix for constraint definitions and naturalness justifications):

(10) WFV as an opaque, Duke-of-York derivation

Stratum 1:

/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:

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

The tableaux above mirror the sound changes in (9). In Stratum 1, the constraints *Coda and *VTV motivate vowel insertion and intervocalic voicing, respectively. In Stratum 2, the constraint *V# motivates word-final vowel deletion, since it's ranked above MAX in this stratum. The result is a vowel that wasn't in the UR, and whose only trace in the SR is the voicing of the word-final stop (i.e. a Duke-of-York derivation similar to the one in Tiberian Hebrew discussed previously).

The above strata demonstrate that WFV *is* predicted by Stratal OT, even if it's limited to natural constraints. If this is learnable, it would represent overprediction by natural constraints, since WFV is an unattested pattern, according to Kiparsky (2006).

3.4. Learnability of opaque Word-final Voicing

In order to test the learnability of the opaque WFV presented in §3.3, I ran simulations similar to those described in §3.2. Using Nazarov and Pater's (to appear) MaxEnt learner, I ran 100 learning simulations with the constraints shown in §3.3 and learning data that demonstrated the UR-SR mapping /kat/ → [kad]. All 100 of these successfully learned WFV, giving a probability of over 97% (M = 0.97918, SD = .0000009) to the correct UR-SR mapping. An example of constraint weights from one of these runs is given in (11).

[5] Technically, this creates a language with voicing intervocalically *and* word-finally. I'll ignore the former environment, since (to my knowledge) having voicing happen word-finally at all is typologically unattested.

(11) Example constraint weights for WFV

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 weights above are analogous to the rankings in (10). The high weights given to *Coda and *VTV in Stratum 1 make vowel insertion and intervocalic voicing highly likely. The second stratum's high weight for *V# makes word-final vowel deletion very likely, creating a pattern that appears to be WFV. These results suggest that WFV—like PND—is predicted by natural constraints in a stratal framework.

4. Discussion

In §3, I explored the predictions that natural constraints make in a stratal framework. I used underprediction and overprediction as metrics for whether a naturalness criterion for constraints is sufficient for restricting typology. §3.1 showed that natural constraints do not underpredict in the case of PND, and §3.2 showed that natural-constraint-based representations of this pattern were learnable. However, §3.3 showed that natural constraints overpredict in the case of WFV, and §3.4 found that this was also reliably learnable. These results suggest that natural constraints are not sufficient, by themselves, for predicting typology. In order to represent an attested, unnatural pattern like PND, Duke-of-York derivations needed to be used. However, these derivations proved too powerful, since they could also represent an unattested, unnatural pattern.

While this isn't meant as an argument for abandoning natural constraints entirely, it does provide evidence that they are not perfectly sufficient for the task of typological prediction.

Appendix

I. Constraint descriptions

A description for each of the constraints used in §3.1 and §3.3 is given below:

Constraint	Description
*[+cont.]/N_	Assign one * for every continuant obstruent in the output that occurs after a nasal.
*[+voice,-cont.]	Assign one * for every voiced stop in the output.
*[+cont.]	Assign one * for every continuant obstruent.
*Coda	Assign one * for every coda consonant.
*VTV	Assign one * for every voiceless obstruent that occurs intervocalically.
*V#	Assign one * for every word-final vowel.
DEP	Assign one * for every segment in the output that does not have a corresponding segment in the input.
MAX	Assign one * for every segment in the input that does not have a corresponding segment in the output.
Faith(F)	Assign one * for every segment in the input that has a different value for feature F in the output.

II. Constraint naturalness justifications

A justification for why each of the markedness constraints used in §3.1 and §3.3 were considered natural is given below:

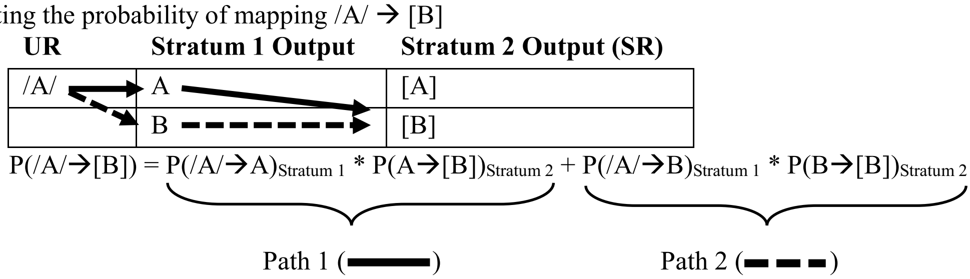
Constraint	**Justification**
*[+cont.]/N_	This constraint could be motivated by the difficulty in producing fricatives after nasals (Vaux 1998) and the fact that prenasalization is not as common for fricatives as it is for stops (Steriade 1993). Zulu, Greek, and Basque all have restrictions on post-nasal fricatives that could be motivated by this constraint (Mielke 2008).
*[+voice,-cont.]	"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).
*[+cont.]	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.
*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.

III. Description of Nazarov and Pater's (to appear) MaxEnt Stratal learner

The input for Nazarov and Pater's (to appear) MaxEnt learner consists of tableaux, a table of learning data, and the number of strata the learner is supposed to use. The tableaux give constraint violations for every relevant input-output mapping in every stratum. The learning data is a table of all the relevant UR→SR mappings and what their probability is in the language being learned. For the simulations in this paper, I used a probability of 1 for the correct UR-SR mappings (i.e. those that demonstrated either PND or WFV).

When the learner acquires a language, it maximizes the likelihood of the learning data by manipulating the constraint weights. It performs this maximization by minimizing the KL-divergence (using the L-BFGS-B method of optimization; Byrd et al. 1995) between the probabilities predicted by the learner's grammar and the probabilities given in the learning data. The predicted probability of a given UR→SR mapping is the sum of the probabilities of every possible path that is consistent with that UR-SR pair. The probability for each path is the product of the probabilities of that path's steps at each stratum. This is illustrated in (a) for the UR-SR mapping of /A/→[B]:

(a) Calculating the probability of mapping /A/ → [B]



The figure and equation in (a) show how this hypothetical predicted probability would be calculated by the learner. Since there are two possible paths to take between the UR /A/ and the SR [B], each path's probability must be summed. Path 1 represents the product of the probability of /A/ being mapped to A, according to Stratum 1's constraint weights and the probability of A being mapped to [B], according to Stratum 2's constraint weights. Likewise, Path 2 is the product of the probabilities of /A/ → B in Stratum 1 and B → [B] in Stratum 2.

The constraint weights determine the probability of a particular mapping at a given stratum. The probability of a mapping given a particular stratum is equal to that mapping's harmony score for the stratum, divided by the sum of all the scores for the possible mappings involving the relevant input in that stratum. The harmony score of a mapping is the sum of all the weighted constraint violations (i.e. the number of violations for a particular constraint, multiplied by that constraint's weight) for a particular input-output mapping at a particular stratum. This is shown below (see Goldwater and Johnson 2003 and Hayes and Wilson 2008 for more on this):

(b) Probability and harmony of an input→output mapping, given a particular stratum

$$H(input \rightarrow output) = \sum_{C} c_i(input \rightarrow output)\, w_i$$

Where:

- *C* is the set of constraints—each of which returns the number of violations for the particular input→output mapping of interest in the relevant stratum,
- *H* is the harmony score for the input→output mapping of interest in the relevant stratum,
- And w is the weight of the constraint of interest in the relevant stratum.

$$p(input \rightarrow output) = \frac{e^{H(input \rightarrow output)}}{\sum_{O} e^{H(input \rightarrow o)}}$$

Where:

- O is the set of all possible outputs for the relevant input.

When the learner has finished optimizing the constraint weights for each stratum, it outputs the final grammar. The weighted constraints then represent a grammar that can give probabilities for any of the relevant UR-SR mappings. For more on how this MaxEnt Stratal learner works, see Nazarov and Pater (to appear), for more information on non-parallel MaxEnt learning, see Staubs and Pater (2016), and see Goldwater and Johnson (2003) and Hayes and Wilson (2008) for more on MaxEnt learners generally.

References

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

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

Byrd, Richard H., Peihuang Lu, Jorge Nocedal, and Ciyou Zhu. 1995. A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM Journal on Scientific Computing 16 (5)*, 1190–1208. doi:10.1137/0916069.

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

Goldwater, Sharon, and Mark Johnson. 2003. Learning OT Constraint Rankings Using a Maximum Entropy Model. In *Proceedings of the Stockholm Workshop on Variation within Optimality Theory,* edited by Jennifer Spenader, Anders Eriksson, and Osten Dahl, 111–20.

Hayes, Bruce, and Colin Wilson. 2008. A Maximum Entropy Model of Phonotactics and Phonotactic Learning. *Linguistic Inquiry 39*, 379–440.

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*, edited by Elizabeth Hume and Keith Johnson, 141–186. San Diego: Academic Press.

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

McCarthy, John (2003). Sympathy, cumulativity, and the Duke-of-York gambit. *The syllable in optimality theory*, edited by Caroline Féry and Ruben van de Vijver, 23-76. Cambridge: Cambridge University Press.

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

Moreton, Elliott, Joe Pater, and Katya Pertsova (2015). Phonological concept learning. *Cognitive science, 41(1)*, 4-69.

Nazarov, Aleksei (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 (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, Martha, and James Amerman (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 *Optimality Theory in Phonology: A Reader*, edited by John J. McCarthy, 271-89. Malden: Blackwell.

Prince, Alan, and Paul Smolensky (2004 [1993]). *Optimality theory: Constraint interaction in generative grammar*. Technical report TR-2, Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, NJ. Technical report CU-CS-696-93, Department of Computer Science, University of Colorado, Boulder. Revised version, 2002: ROA-537-0802, Rutgers Optimality Archive, http://roa.rutgers.edu. Published 2004, Oxford: Blackwell.

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

Staubs, Robert, and Joe Pater (2016). Learning Serial Constraint-Based Grammars. In *Harmonic Grammar and Harmonic Serialism*, edited by John McCarthy and Joe Pater, 1-16. London: Equinox Press. Available at: http://people.umass.edu/pater/staubs-pater-serial-learning.pdf

Tesar, Bruce (1998). An iterative strategy for language learning. *Lingua*, *104(1-2)*, 131-145.

Wilson, Colin (2006). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive science, 30(5)*, 945-982.

Proceedings of the 35th West Coast Conference on Formal Linguistics

edited by Wm. G. Bennett, Lindsay Hracs, and Dennis Ryan Storoshenko

Cascadilla Proceedings Project Somerville, MA 2018

Copyright information

ISBN 978-1-57473-472-0 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Prickett, Brandon. 2018. Post-nasal Devoicing as Opacity: A Problem for Natural Constraints. In *Proceedings of the 35th West Coast Conference on Formal Linguistics*, ed. Wm. G. Bennett et al., 331-340. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #3405.