Some Formal Implications of Deletion Saltation

Jennifer L. Smith

UNC Chapel Hill

Version date: January 29, 2021

To appear in *LI*

In a phonological *saltation* alternation, a segment or class “skips” a relatively similar category to surface as something less similar, as when /g/ alternates with [x], skipping [k]. White (2013) and Hayes and White (2015) argue that saltation is *unnatural*—difficult to learn in the laboratory and diachronically unstable. They propose that the phonological grammar includes a *learning bias* against such unnatural patterns. White and Hayes further demonstrate that Harmonic Grammar (HG; Legendre, Miyata, and Smolensky 1990) cannot model typical saltation without nondefault mechanisms that would require extra steps in acquisition, making HG consistent with their proposed learning bias.

I identify *deletion saltation* as a distinct saltation subtype and show that HG, with faithfulness formalized in standard Correspondence Theory (CT; McCarthy and Prince 1995), *can* model this pattern. HG/CT thus predicts that deletion saltation, unlike typical (here called *segment-scale*) saltation, is natural. Other frameworks fail to distinguish the two saltation types—they can either model both types, or neither. Consequently, if future empirical work finds deletion saltation to be more natural than other saltation patterns, this would support weighted-constraint models such as HG over ranked-constraint models such as Optimality Theory (OT; Prince and Smolensky 1993, 2004); would support CT over the *Map model of faithfulness (Zuraw 2013); and would support formalizing CT featural-faithfulness constraints in terms of Ident constraints, binary features, or both.

1 Two Kinds of Saltation

White (2013) and Hayes and White (2015) use the term *saltation* for alternations of the following type: underlying /A/ maps to [C] (/A/→[C]), “skipping” potential outcome **[B]* (/A/→**[B]*), even though **[B]* is intermediate between /A/ and [C]: both /A/ and

*_[B]_, and *_[B]_ and [C], are more similar than /A/ and [C] are, and *_[B]_ shares all properties that /A/ and [C] have in common.[1] Moreover, /B/ surfaces faithfully (/B/→[B], not /B/→*[C]), so saltation is a case of phonological opacity—the *_[B]_→[C] mapping subsumed in /A/→[C] is non-surface-true (McCarthy 1999). I schematize this pattern /A/–*_[B]_–[C], and I label it _segment-scale saltation,_ because White (2013:19) and Hayes and White (2015:267) explicitly define /A/, *_[B]_, and [C] as segments.

An example from Colloquial Northern German is shown in (1). Coda /g/, which must devoice, skips *_[k]_ and maps to [x], even though /g/ and [x] differ by [±continuant] and [±voice], while *_[k]_ differs from /g/ only by [±voi]. Coda /k/, however, maps to [k].

(1) Segment-scale saltation: /g/–*_[k]_ –[x] in codas (Ito and Mester 2003:274, 291)

a.	/g/	/tsoː**g**/	→ [tsoː**x**], *[tsoː**k**]	_cf._	[tsoː**g**-ən]	‘pulled’, 1sg./1pl.
		/tʀuː**g**/	→ [tʀuː**x**], *[tʀuː**k**]		[tʀuː**g**-ən]	‘carried’, 1sg./1pl.
		/flyː**g**/	→ [fluː**x**], *[fluː**k**]		[flyː**g**-ə]	‘flight’, sg./pl.
b.	_vs._ /k/	/dɪ**k**/	→ [dɪ**k**], *[dɪ**x**]		[dɪ**k**-ə]	‘thick’, pred./attrib.pl.

Alternations as in (1) have been called _phonological derived-environment effects_ (PDEEs; Łubowicz 2002, following Kiparsky 1973, 1982, 1993; Mascaró 1976; Rubach 1984) because, in derivational terms, some process (here, spirantization) applies only if its target ([k]) is derived by an earlier process (/g/ devoicing). White and Hayes prefer the term _saltation_ (Minkova 1993, Lass 1997) on the grounds that phonological derivations are irrelevant in parallel HG or OT. However, their definition of (segment-scale) saltation is not synonymous with PDEE, because there are also PDEEs where a segment “skips” something to alternate with zero. This is the pattern that I call _deletion saltation_.

Deletion saltation can be schematized /A/–**[B]*–Ø: /A/ deletes entirely rather than changing to **[B]*, although [B] from /B/ survives. In an example from Standard German, coda /g/ following [ŋ] skips **[k]* and deletes (2a), while underlying /ŋk/ survives (2b), and voiced obstruent codas are avoided elsewhere by devoicing instead of deletion (2c).

(2) Deletion saltation: /g/–**[k]*–Ø in coda /[ŋ]_ (Ito and Mester 2003:274, 276, 289)

a. /dɪftɔŋ**g**/ → [dɪftɔŋ_], *[dɪftɔŋ**k**] ‘diphthong’ *cf.* [dɪf.tɔŋ.**g**iː.ʀən] ‘to diphthongise’

b. *vs.* /baŋ**k**/ → [baŋ**k**], *[baŋ_], ‘bank’

c. *vs.* /taː**g**/ → [taː**k**], *[taː_] ‘day’ *cf.* [taː.**g**ə] ‘days’

/han**d**/ → [han**t**], *[han_] ‘hand’ [hɛn.**d**ə] ‘hands’

Formally, the deletion pattern /A/–**[B]*–Ø is also a PDEE. In derivational terms, one process—specifically, deletion—applies only if its target (here, [k]) is derived by an earlier process (/g/ devoicing). Section 5 gives a nonderivational formal definition of PDEEs, as *triggering-type cumulative M&F interactions*, which likewise includes both saltation types. Conceptually, deletion saltation still involves “skipping”, on the view that deleting a segment is more extreme than altering features (see, e.g., Lass 1984:178 on lenition, where deletion is included as the endpoint of a trajectory of feature changes).

2 Only Deletion Saltation Can Be Modeled in HG/CT

The deletion saltation pattern in (2) can be analyzed in HG under Correspondence Theory (CT; McCarthy and Prince 1995), as seen in (3). Here and in the HG tableaus that follow, the number of constraint violations in each cell is shown as a negative integer, and a representative weight for each constraint (compatible with its general weighting conditions) is given at the top of its column. The harmony of a candidate (ℋ) is the sum of its weighted violations, and the candidate with the highest harmony—closest to 0—wins.

(3) Deletion saltation in HG/CT: /ɡ/–*[k] – Ø

/dɪftɔŋɡ/	*VoiObstCoda	Max	*VelarNC	Ident[±voi]	ℋ
Weights	*10*	*4*	*3*	*2*	
→ a. dɪf.tɔŋ_		–1			*–4*
b. **dɪf.tɔŋk*			–1	–1	*–5*

In (3), the saltation candidate (a) violates Max, because /ɡ/ has been deleted. The “skipped” competitor (b) violates *VelarNC, because it has a velar nasal+C coda, and Ident[±voi], because /ɡ/ has been devoiced. (Only candidates satisfying high-weighted *VoiObstCoda are shown.) Max outweighs *VelarNC and Ident[±voi], because deletion is not generally chosen to avoid velar NC codas (2b) or devoicing (2c). But, crucially, deletion is preferred to a devoiced stop *in* a velar NC coda cluster. If the *combined* weight of *VelarNC and Ident[±voi] is greater than that of Max, as in (3), then (a) can win.

This kind of *cumulative constraint interaction*, also called a *gang effect*, is a key characteristic of HG, since constraints are weighted, rather than strictly ranked as in OT. Pater (2009, 2016) demonstrates that gang effects arise only when constraint violations include what he calls an *asymmetric trade-off* (ATO). This result is important, because it means that gang effects in HG are actually rather restricted. Informally, an ATO occurs under the following conditions (see Pater 2009, 2016 for a more rigorous description). First, some *competitor* does better than the *winner* on a higher-weighted constraint, as in (3), where competitor (b) outperforms winner (a) on Max. In addition, the competitor has a greater number of *unshared* violations of lower-weighted constraints. In (3), (b) has violations of

*VelarNC and Ident[±voi] that (a) does not share. This scenario is called an asymmetric trade-off because, in the competition between the two critical candidates, one violation of Max for (a) is *traded against* a total of two violations of *VelarNC and Ident[±voi] for (b)—a violation-count ratio that is is *asymmetric* (not 1:1).

It is the lack of an ATO that makes segment-scale saltation incompatible with HG (without some additional mechanism). White (2013:sec. 2.4.3), discussing Campidanian Sardinian, shows that the saltation candidate cannot win under HG/CT; here, (4), attempting (1), makes the reason explicit. With no ATO, there can be no gang effect.

(4) Unsuccessful HG/CT analysis of segment-scale saltation: intended /g/–_*[k]_ –[x]

/tsoːg/	*VoiObstCoda	Ident[±cont]	*DorsPlos	Ident[±voi]	ℋ
Weights	*10*	*4*	*3*	*n*	
(→) a. tsoːx		$\boxed{-1}$		–1 | *shared*	*–(4+n)*
✗ b. _*tsoːk_			$\boxed{-1}$	–1 | *shared*	*–(3+n)*

The intended winner (see (1)) is the saltation candidate (a), where the voiced stop /g/ surfaces as a voiceless fricative [x], violating both Ident[±cont] and Ident[±voi]. The “skipped” competitor (b), with voiceless stop _*[k]_, likewise violates Ident[±voi], and also *DorsalPlosive. But Ident[±cont] outweighs *DorsPlos, because /k/→[k], not *[x] (1b). This means that (a) could only win if *DorsPlos were involved in a cumulative interaction. The problem for (a) is that shared violations do not contribute to gang effects (Pater 2009, 2016). The Ident[±voi] violation is shared, so any contribution by this constraint to lowering (b)’s harmony (weight=*n*) penalizes (a) to the same extent. The only trade-off in (4) is *symmetric*, as shown by the boxed violations: one Ident[±cont] violation for (a) trades against one *DorsPlos violation for (b). There is no gang effect, so (a) cannot win.

The lack of an ATO to rescue the saltation candidate is a general result for segment-scale saltation. Given /A/–**[B]* –[C] where **[B]* is featurally intermediate between /A/ and [C], /A/ must differ from both **[B]* and [C] by the same feature [±*f*], which means that the mappings /A/→[C] and /A/→**[B]* necessarily share an Ident[±*f*] violation. This violation, being shared, never contributes to a gang effect favoring the saltation candidate. As a result, and as desired according to White (2013; Hayes and White 2015), HG/CT cannot generate segment-scale saltation without some additional formal mechanism.

What is different about the deletion saltation case in (3) is that the Ident[±voi] violation is not shared between saltation candidate (a) and its competitor (b). Ident[±*f*] is only violated when corresponding (input and output) segments differ in [±*f*] (McCarthy and Prince 1995). Deletion allows (a) to avoid the Ident[±voi] violation incurred by (b), since (a) has no correspondent for the input /g/. This lets Ident[±voi] contribute to an ATO: one Max violation by (a) trades against two violations, of *VelarNC and Ident[±voi], by (b). If *w*(Max) < (*w*(*VelarNC) + *w*(Ident[±voi]))—a gang effect—then (a) wins.

Thus, HG/CT predicts that only deletion saltation should be natural.

3 Ranked Constraints Fail to Distinguish the Saltation Types

Unlike HG, with its weighted constraints, ranked-constraint frameworks are unable to distinguish segment-scale saltation from deletion saltation. Classic OT (Prince and Smolensky 1993, 2004), in which constraints are never cumulative, is intrinsically unable to model any cumulative interaction, with or without an ATO, and so necessarily fails to differentiate the saltation types. But even OT models with additional formalisms designed to allow cumulative interaction, either directly as in local constraint conjunction (LCC;

Łubowicz 2002; Ito and Mester 2003) or indirectly as in comparative markedness (McCarthy 2003), are unable to distinguish between segment-scale and deletion saltation.

For example, an OT-LCC approach to the pattern in (1) could introduce a *conjoined constraint* *DorsPlos&*Seg*Ident[±voi], which is violated only if both component constraints are violated *locally*, in this case, within the same segment (as specified by the local-conjunction operator &*Seg*).[2] Ranking this constraint above Ident[±cont] in an OT counterpart to tableau (4) would allow saltation candidate (a) to win. Deletion saltation, as in (2), could be handled in exactly the same way, with a conjoined constraint *VelarNC&*Seg*Ident[±voi] that outranks Max in an OT counterpart to (3). A conjoined constraint assigns violations separately from its component constraints, so it is irrelevant whether the cumulative constraint interaction involves a shared violation (as with *DorsPlos&*Seg*Ident[±voi], for segment-scale saltation), or not (as with *VelarNC&*Seg* Ident[±voi], for deletion saltation). As a result, OT-LCC treats segment-scale and deletion saltation as equally possible, predicting no difference in naturalness between the two.

Comparative markedness (McCarthy 2003) likewise fails to distinguish the two saltation types; the "new markedness" constraints in this framework give rise to cumulative effects whether or not constraint violations are shared between candidates.

4 Implications for Faithfulness Constraints and Phonological Features

Under HG, the crucial difference between deletion saltation and segment-scale saltation is that only deletion saltation involves an ATO. But this itself depends on the way that phonological features, and the faithfulness constraints governing them, are formalized. If *Map constraints (Zuraw 2007, 2013) are used to model faithfulness, no formal distinction between the two saltation types is predicted (section 4.1). Even under CT, deletion saltation results in an ATO only if there are IDENT[(±)*f*] constraints, or features are binary, or both (section 4.2); otherwise, the two saltation types are formally identical.

*4.1 *MAP Constraints Fail to Distinguish the Two Saltation Types*

White (2013, 2017) and Hayes and White (2015) address a key puzzle about (segment-scale) saltation: why is it unnatural, but still attested? They propose that there is a defeasible learning bias against saltation—the bias is what makes saltation unnatural, but the ability to override it, given appropriate learning data, is why saltation exists.

To implement this anti-saltation bias, White and Hayes adopt the *MAP model of faithfulness (Zuraw 2007, 2013). Each *MAP constraint penalizes, not a change in a feature value, but a particular mapping from one segment or class to another: *MAP(g,k), penalizing a change [g]→[k], is formally independent of *MAP(g,x), even though both mappings are unfaithful to voicing. The anti-saltation bias emerges because *MAP constraints form a default hierarchy in which perceptually more-salient changes (based on the P-map; Steriade (2008)) take priority: *w*(*MAP(g,x)) > *w*(*MAP(g,k)). In the /g/–**[k]*–[x] pattern in (1), the saltation candidate (a) violates higher-priority *MAP(g,x)), while the competitor (b) violates only *MAP(g,k). The saltation candidate thus cannot win under the default *MAP hierarchy. White and Hayes propose that the acquisition of a segment-scale saltation pattern requires the learner to take an additional step, overriding the default *MAP ranking and promoting *MAP(g,k) over *MAP(g,x). The burden imposed by this additional step is the formal implementation of the learning bias.

Abandoning CT for *MAP raises various questions,[3] but of greatest relevance here, the

*Map approach does not systematically predict a difference between the two saltation types. Under *Map, the /g/–*[k]–Ø deletion-saltation case in (1) would only be “natural” if a change from [g] to Ø were *less* perceptually salient than a change from [g] to [k]. Confusion matrices in white noise (Phatak, Lovitt, and Allen 2008) actually show the contrary: [g] is misperceived as [k] *more* often than as Ø (until a signal-to-noise ratio of –15dB or worse). While the relative salience of /A/→*[B] and /A/→Ø might vary for different values of /A/ and *[B], it is clear that the *Map approach does not predict a systematic difference in naturalness between segment-scale and deletion saltation.

4.2 Implications for CT Faithfulness Constraints and Distinctive Features

In the HG/CT analyses of saltation patterns presented in (3) and (4) above, featural faithfulness is mediated indirectly (through segments), by Ident constraints. Features themselves are binary, [±*f*]. Even under CT, deletion saltation gives rise to an ATO in HG only if at least one of these assumptions is maintained.

An alternative version of CT has features in direct correspondence, subject to Max (anti-deletion) and Dep (anti-insertion) constraints (McCarthy and Prince 1995). If Ident is replaced by Max and Dep, but features are binary [±*f*], an ATO arises in deletion saltation, as in (5) (where Max-*other* encapsulates Max-*seg* and all Max[*f*] other than Max[+voi]). Saltation candidate (a) and competitor (b) share a violation of Max[+voi], which (unlike Ident[±voi]) is violated when a segment deletes. Nevertheless, faithfulness still distinguishes deletion (a) from devoicing (b): only devoicing violates Dep[–voi]. Two unshared violations for (b) trade against one for (a), and this ATO allows (a) to win.[4]

(5) Deletion saltation: ATO with Max/Dep and binary [±*f*]

/dɪftɔŋɡ/	*NG#	Max-*other*	*VelarNC	Dep[–voi]	Max[+voi]	ℋ
Weights	*10*	*4*	*3*	*2*	*n*	
→ a. dɪftɔŋ_		$\boxed{-1}$			–1 | *shared*	*–(4+n)*
b. **dɪftɔŋk*			$\boxed{-1}$	$\boxed{-1}$	–1 | *shared*	*–(5+n)*

An alternative to binary features ([+*f*] or [–*f*]) is for some or all features to be privative—present ([*f*]) or absent (e.g., Lombardi 1994; Steriade 1995). Even with privative features, Ident constraints are unviolated when segments delete, so deletion saltation would maintain the ATO. But in a model with *both* Max/Dep[*f*] *and* privative [*f*], deletion saltation no longer creates an ATO (at least not when the competitor **[B]* undergoes feature deletion as opposed to insertion). The analysis in (6) crucially differs from (5) in that there is no [–voi] feature, so (b) does not violate Dep[–voi].[5]

(6) Deletion saltation: No ATO with Max/Dep[*f*] and privative [*f*], if [*f*] is deleted

/dɪftɔŋɡ/	*NG#	Max-*other*	*VelarNC	Max[voi]	ℋ
Weights	*10*	*4*	*3*	*n*	
(→) a. dɪftɔŋ_		$\boxed{-1}$		–1 | *shared*	*–(4+n)*
✗ b. **dɪftɔŋk*			$\boxed{-1}$	–1 | *shared*	*–(3+n)*

Even in HG, then, a difference in naturalness between deletion saltation and segment-scale saltation is predicted only if featural faithfulness is formalized in terms of Ident[*f*] constraints, binary features, or both.

5 Further Implications: Cumulative M&F Interaction in HG/CT

Deletion saltation in HG/CT involves cumulative interaction between a markedness constraint (M), which evaluates only outputs, and a faithfulness constraint (F), which penalizes input/output differences. The role of such M&F interaction in the grammar has been controversial. Certain patterns are attested, so classic OT's exclusion of *all* cumulative interaction is too strong. But allowing *any* logically possible M&F interaction makes problematic typological predictions (e.g., McCarthy 2003; Łubowicz 2005; Hayes and White 2015). This section considers the implications of HG/CT for general patterns of M&F interaction. Only cases that involve an ATO, like deletion saltation, can be directly modeled in HG/CT and are predicted by this framework to be natural.

M&F interaction is at work in any phenomenon where some marked surface structure is tolerated when it is faithful to an underlying form (violating M only), but crucially not when it results from a phonological process (violating both M and F). Two subclasses can be distinguished (McCarthy 2003): triggering cases and blocking cases.

Triggering M&F interactions are phonological derived-environment effects (PDEEs), including both deletion saltation and segment-scale saltation. In a PDEE, some process, alleviating a markedness violation, is *triggered* in a context where another process also applies, creating a faithfulness violation. In the absence of this concurrent faithfulness violation, the markedness violation is tolerated and the marked structure survives.

For the deletion-saltation pattern in (2), the M&F interaction is between *VelarNC (M) and Ident[±voi] (F); deletion is chosen only if both constraints would otherwise be violated. For the segment-scale pattern in (1), the M&F interaction is between *DorsalPlosive (M) and Ident[±voi] (F); in the intended saltation grammar, spirantization is to be chosen only if both constraints are at stake. As seen above, these M&F violations form an ATO—producing a cumulative effect in HG/CT by default—only for deletion saltation.

The other subclass of M&F interaction, the *blocking* type, is the pattern known as a *grandfather effect* (McCarthy 2003; see also Hall 2006 and the Target Conditions of Archangeli and Pulleyblank 1994). Some process that generally applies, creating a faithfulness violation, is *blocked* if it would result in a marked structure, incurring a simultaneous markedness violation. Just as for a PDEE, the relevant marked structure is tolerated if underlying (when only the markedness violation is assigned). One example from McCarthy (2003) is Mekkan Arabic (Bakalla 1973; Abu-Mansour 1996), where obstruent voicing assimilation is blocked if it would *create* a voiced coda, but voicing assimilation and voiced obstruent codas are allowed independently. Because grandfather effects involve the blocking of a process (whose application would violate F) precisely where it would create a marked structure (violating M), the relevant M&F interaction should, under HG, systematically produce an ATO.[6] Inspection of the grandfather effects catalogued in McCarthy (2003) confirms this for each example discussed there.[7]

What about pathological cases of M&F interaction, such as Hayes and White's (2015) hypothetical language where obstruent voicing contrasts surface faithfully only between consonants? Pater (2016) demonstrates that the ATO condition on cumulative constraint interaction in HG eliminates many of the pathological cumulativity patterns that can be modeled with LCC and other OT-based approaches. Concretely, the problematic cases of M&F interaction under OT-LCC constructed by Hayes and White (2015:284) and McCarthy (2003:27) both involve shared violations between the plausible and pathological candidates, so there is no ATO and no gang effect under HG.

In summary, among cumulative M&F interaction patterns, only those that give rise to an ATO can be modeled in HG/CT without recourse to additional formal mechanisms, and are therefore predicted to be phonologically natural patterns. In addition to deletion saltation, this includes the class of grandfather effects. Excluded from the class of patterns predicted to be natural is segment-scale saltation (which is as desired, following White and Hayes), as well as many of the hypothetical patterns of concern under OT-LCC.[8]

6 Conclusions

HG can model only deletion saltation, not segment-scale saltation (without nondefault mechanisms). By contrast, the two saltation types are equivalent in OT—possible with a supplementary formalism such as LCC, and impossible otherwise. This result establishes a novel difference in empirical predictions between the two frameworks: if deletion saltation turns out to be more "natural" (more learnable or diachronically stable) than segment-scale saltation, this finding would support HG over OT.

There are implications for featural faithfulness as well. First, a difference between the saltation types is predicted under Correspondence Theory, but not under *Map. Second, only certain implementations of featural faithfulness predict a difference between deletion and feature-scale saltation: either the grammar must include Ident[±*f*] constraints; or, if featural faithfulness is mediated by Max and Dep, then features must be binary.

If deletion saltation does prove to be more natural than segment-scale saltation, the key insight about an anti-saltation learning bias from White (2013, 2017) and Hayes and White (2015) can still be implemented in HG/CT even without *Map. HG/CT is, as needed, a grammar framework whose default state disallows segment-scale saltation and allows deletion saltation. Along with this, we would need a special mechanism that a learner can invoke if segment-scale saltation is present in the ambient data. Perhaps local constraint conjunction could be formalized as an option that is parochially recruited for a specific constraint pair when needed, in the same spirit as White's and Hayes's proposal for a language-specific reranking of *MAP constraints.[9]

Both saltation patterns are relatively rare. Hayes and White (2015) list all cases of segment-scale saltation of which they are aware, numbering about ten, and the only additional deletion saltation case known to me other than (2) is a pattern of /a,e/–*[i]–Ø in Modern Hebrew that is to some extent morphologically conditioned (Bat-El 2008). Nevertheless, these patterns potentially have an important theoretical contribution to make. The next step is to determine whether deletion saltation does, in fact, turn out to be more natural than segment-scale saltation.

References

Abu-Mansour, Mahasen Hasan. 1996. Voice as a privative feature. In *Perspectives on Arabic Linguistics VIII*, ed. by Mushira Eid, 201–231. Amsterdam: Benjamins.

Archangeli, Diana, and Douglas Pulleyblank. 1994. *Grounded phonology.* Cambridge, MA: MIT Press.

Bakalla, Muhammad Hasan. 1973. The morphology and phonology of Meccan Arabic. Doctoral dissertation, School of Oriental and African Studies, University of London.

Bat-El, Outi. 2008. Morphologically conditioned V–Ø alternation in Hebrew: Distinction among nouns, adjectives & participles, and verbs. In *Current issues in generative Hebrew linguistics*, ed. by Sharon Armon-Lotem, Gabi Danon, and Susan Rothstein, 27–60. Amsterdam: Benjamins.

Hall, T. Alan. 2006. Derived environment blocking effects in Optimality Theory. *Natural Language and Linguistic Theory* 24:803–856.

Hayes, Bruce, and James White. 2015. Saltation and the P-map. *Phonology* 32:267–302.

Ito, Junko, and Armin Mester. 2003. On the sources of opacity in OT: Coda processes in German. In *The syllable in Optimality Theory*, ed. by Sharon Armon-Lotem, Gabi Danon, and Susan Rothstein, 271–303. Cambridge: Cambridge University Press.

Jesney, Karen. 2011. Cumulative constraint interaction in phonological acquisition and typology. Doctoral dissertation, University of Massachusetts, Amherst.

Jessen, Michael, and Catherine Ringen. 2002. Laryngeal features in German. *Phonology* 19:189–218.

Kiparsky, Paul. 1973. Abstractness, opacity and global rules. In *Three dimensions of linguistic theory*, ed. by Osamu Fujimura, 57–86. Tokyo: TEC.

Kiparsky, Paul. 1982. From cyclic phonology to lexical phonology. In *The structure of phonological representations*, vol. 1, ed. by Harry van der Hulst and Norval Smith, 131–175. Dordrecht: Foris.

Kiparsky, Paul. 1993. Blocking in non-derived environments. In *Studies in Lexical Phonology*, ed. by Sharon Hargus and Ellen Kaisse, 277–313. Phonetics and Phonology 4. San Diego, CA: Academic Press.

Lass, Roger. 1984. *Phonology: An introduction to basic concepts*. Cambridge: Cambridge University Press.

Lass, Roger. 1997. *Historical linguistics and language change*. Cambridge: Cambridge University Press.

Legendre, Géraldine, Yoshiro Miyata, and Paul Smolensky. 1990. Can connectionism contribute to syntax?: Harmonic grammar, with an application. Report CU-CS-485-90. Boulder, CO: University of Colorado at Boulder Department of Computer Science.

Lombardi, Linda. 1994. *Laryngeal features and laryngeal neutralization*. New York: Garland.

Łubowicz, Anna. 2002. Derived environment effects in Optimality Theory. *Lingua* 112:243–280.

Łubowicz, Anna. 2005. Locality of conjunction. In *Proceedings of the 24th West Coast Conference on Formal Linguistics*, ed. by John Alderete, Chung-hye Han, and Alexei Kochetov, 254–262. Somerville, MA: Cascadilla Proceedings Project.

Mascaró, Joan. 1976. *Catalan phonology and the phonological cycle*. Doctoral dissertation, MIT, Cambridge, MA.

McCarthy, John. 1999. Sympathy and phonological opacity. *Phonology* 16:331–399.

McCarthy, John. 2003. Comparative markedness. *Theoretical Linguistics* 29:1–51.

McCarthy, John, and Alan Prince. 1995. Faithfulness and reduplicative identity. In *Papers in Optimality Theory*, ed. by Jill Beckman, Suzanne Urbanczyk, and Laura Walsh Dickey, 249–384. University of Massachusetts Occasional Papers 18. Amherst, MA: GLSA.

Minkova, Donka. 1993. On leapfrogging in historical phonology. In *Historical Linguistics 1991*, ed. by Jaap van Marle, 211–228. Amsterdam: Benjamins.

Pater, Joe. 2007. The locus of exceptionality: Morpheme-specific phonology as constraint indexation. In *Papers in Optimality Theory III*, ed. by Leah Bateman, Michael O'Keefe, Ehren Reilly and Adam Werle, 259–296. University of Massachusetts Occasional Papers 32. Amherst, MA: GLSA.

Pater, Joe. 2009. Weighted constraints in generative linguistics. *Cognitive Science* 33:999–1035.

Pater, Joe. 2016. Universal Grammar with weighted constraints. In *Harmonic Grammar and Harmonic Serialism*, ed. by John McCarthy and Joe Pater, 1–46. London: Equinox.

Phatak, Sandeep A., Andrew Lovitt, and Jont B. Allen. 2008. Consonant confusions in white noise. *Journal of the Acoustical Society of America* 124:1220–1233.

Prince, Alan, and Paul Smolensky. 1993. Optimality Theory: Constraint interaction in generative grammar. RuCCS Technical Report. Piscataway, NJ: Rutgers University Center for Cognitive Science.

Prince, Alan, and Paul Smolensky. 2004. *Optimality Theory: Constraint interaction in generative grammar*. Oxford: Blackwell.

Rubach, Jerzy. 1984. *Cyclic and Lexical Phonology: The structure of Polish.* Dordrecht: Foris.

Smith, Caitlin, and Charlie O'Hara. 2020. Learnability of derivationally opaque processes in the Gestural Harmony Model. Ms., Johns Hopkins University and University of Southern California. Baltimore, MD, and Los Angeles.

Steriade, Donca. 1995. Underspecification and markedness. In *The handbook of phonological theory*, ed. by John Goldsmith, 114–74. Oxford: Blackwell.

Steriade, Donca. 2008. The phonology of perceptibility effects: The P-map and its consequences for constraint organization. In *The nature of the word: studies in honor of Paul Kiparsky*, ed. by Kristin Hanson and Sharon Inkelas, 151–180. Cambridge, MA: MIT Press.

White, James. 2013. *Bias in phonological learning: Evidence from saltation.* Doctoral dissertation, University of California, Los Angeles.

White, James. 2017. Accounting for the learnability of saltation in phonological theory: A maximum entropy model with a P-map bias. *Language* 93:1–36.

Zuraw, Kie. 2007. The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog. *Language* 83:277–316.

Zuraw, Kie. 2013. *MAP constraints. Ms., University of California, Los Angeles. Available at http://linguistics.ucla.edu/people/zuraw/dnldpprs/star_map.pdf.

309 Smith Building, CB #3155
University of North Carolina
Chapel Hill, NC 27510 USA
jlsmith@email.unc.edu

Notes

Thanks to Michael Becker, Bruce Hayes, Brian Hsu, Armin Mester, Elliott Moreton, Katya Pertsova, Caitlin Smith, Jamie White, audience members at LSA 2018, and two anonymous reviewers for comments and discussion, and to the Schwab Academic Excellence Fund, Institute for the Arts and Humanities, UNC Chapel Hill, for research support.

[1] White (2013:sec.2.1) frames the /A/–**[B]*–[C] scale in general terms of *phonetic similarity* (*phonologically relevant similarity* would be more precise). He notes that such a scale could be modeled with features, and Hayes and White (2015:267) consider this. But White (2013) and Hayes and White (2015) ultimately implement a perceptual-similarity scale, because their analysis uses *Map faithfulness constraints. In section 4, I show that a formal distinction between saltation types is only captured with a featural approach to faithfulness, so I model similarity accordingly, as shared feature values.

[2] Ito and Mester's (2003) OT-LCC analyses of the saltation cases in (1) and (2) use slightly different constraint formalizations than are given here.

[3] For example, Zuraw's (2013) proposal allows *Map constraints only on output-output correspondence, on the grounds that inputs have no phonetic salience. White and Hayes argue that saltation can be handled with output-output faithfulness. But in general, adopting the *Map model requires formalizing faithfulness for correspondence relations involving the input differently from other relations (output-output, base-reduplicant, etc.).

[4] Since Max-*other* encapsulates multiple constraints, this is not technically "one" violation. But this scenario is still an ATO; weights can be assigned so that the violations included in Max-*other* outweigh *VelarNC and Dep[–voi] separately, but the latter together

outweigh Max-*other*. See Pater (2016) for more on the formal definition of an ATO.

[5] The tableau in (6) assumes that [g] has [voi] and [k] is unspecified. Jessen and Ringen (2002) argue that the Standard German contrast is actually between *plain* /g/ and *aspirated* /k/; if so, then the privative-feature analysis for this alternation would involve the insertion of [spread glottis] rather than the deletion of [voi]. The point of (6) is to illustrate the general result that, if there is feature deletion in the competitor candidate, then there is no ATO in deletion saltation, assuming privative [*f*] and Max[*f*]/Dep[*f*].

[6] McCarthy (2003) emphasizes that the triggering and blocking types—PDEEs and grandfather effects—are formally identical under OT with comparative markedness (or LCC): in both types, the loser fatally incurs a cumulative M&F (or "new markedness") violation. This differs from HG/CT, where the key role of ATOs in cumulative interaction makes grandfather effects and deletion saltation distinct from segment-scale saltation.

[7] Jesney (2011) gives an HG analysis for Mekkan Arabic using a different set of constraints. Her approach still involves an ATO, but it is one with M&*M* interaction.

[8] As for *morphological* DEEs, Pater (2007) analyzes them in HG with morpheme-specific indexed constraints, requiring no cumulative M&F interaction. MDEEs are thus unrelated to PDEEs in HG, despite their similarity in derivational frameworks.

[9] See also C. Smith and O'Hara (2020) for another approach: a learning bias against saltation that is implemented computationally.