Adam Albright (MIT)—albright@mit.edu

Cumulative violations and complexity thresholds: Evidence from Lakhota

1 Introduction

(1) A design feature of phonology?

- Phonological repairs appear to be largely independent of one another
- Would be violated by the following:
  - Final devoicing *unless* (or *only if*) nasal place assimilation has also occurred
  - Cluster simplification if multiple clusters within the word
- Such interactions are typically excluded by limits on how rules/constraints interact
  - Rules: apply if structural description is met, regardless of which other rules apply
  - Constraints: if a markedness violation is repaired (Markedness $\gg$ Faithfulness), the number of violations or presence of other markedness violations is irrelevant (OT: Prince and Smolensky 1997; Harmonic Grammar: Pater, in press)
  - Allowing repair of one violation does not depend on presence of another violation
- A consequence: no complexity thresholds
  - Tough to capture: either final voiced stops or disagreeing clusters, but not both
  - Also tough: one cluster per word, but not two (or 3…*n*)

(2) An important distinction: common vs. uncommon bans on cooccurring elements

- A widely observed restriction: multiple occurrences of the same feature in the word
  - Ruled out by constraints that specifically target such combinations (OCP[feature]: e.g., *C^h…C^h), not double violation of simpler constraints (*C^h)
  - May be related to articulatory (Walter 2007) or perceptual (Gallagher 2009) considerations
  - Typically targets specific features, such as place (Frisch, Pierrehumbert, and Broe 2004; Coetzee and Pater 2008); aspiration & ejection (MacEachern 1999; Rose and Walker 2004)
  - Supports the idea that attested 'one but no more' restrictions involve special constraints on multiple occurrences of particular features
- Not widely observed
  - Ban on other repeated features/structures: *S…S (two fricatives), or *CC…CC (two clusters)
  - Ban on cooccurrence of "independent" elements: ejectives and clusters cannot cooccur
    - ☞ Would require constraints like *C'…CC or *CC…C'
  - Such constraints could be constructed with a unrestricted form of constraint conjunction (Smolensky 1995; ; cf. Moreton and Smolensky 2002)
  - Plausibly dispreferred or excluded (complex, no known phonetic motivation)

(3) Goals of this talk

- Document several unexpected cooccurrence restrictions in Lakhota
  - *S…S (more than one fricative); *CC…CC (more than one cluster)
  - *S…CC, *CC…S: combinations of clusters and fricatives
  - *C'…CC, *CC…C': combinations of clusters and ejective/aspirated stops
  - These restrictions require more complex constraint interactions than standard OT/Harmonic Grammar allow
- However, they also appear to obey some important limitations
  - Lexically gradient: cooccurrences are underattested, but not banned outright
  - Morpheme structure conditions: not enforced in morphologically complex words
  - Target combinations of elements which are independently relatively rare in the language
  - Conjoined constraints (*CC and *C') do not predict or require these limitations
- Proposal: additive interactions of weighted constraints
  - "Ganging up": violations of independent markedness constraints (*CC, *C', etc.) may sum to be worse than either violations alone
  - Severity of multiple violations follows from severity of independent violations
  - Difficult to ban combinations completely while allowing independent parts
  - May emerge as a gradient effect, particularly if independent parts are already dispreferred

2 Cooccurrence restrictions in Lakhota

(4) Background on Lakhota

- Siouan language, spoken primarily in South Dakota (Pine Ridge, Rosebud reservations)
- ≈6,000–10,000 speakers (all levels) as of 2000 census
- Data sources: Munro (1989), Buechel and Manhart (2002), field notes (2000–2002) from work with Mary Rose Iron Teeth

(5) Lakhota Consonant inventory

- Stops and affricates: three-way contrast

		Initial		Medial	
Voiceless unaspirated	p, t, tʃ, k	**k**a	'there'	ʃ'a**k**a	'strong'
Aspirated	p^h, t^h, tʃ^h, k^h	**k^h**a	'to mean'	mã**k^h**a	'earth'
Ejective	p', t', tʃ', k'	**k'**a	'to dig'	t͡ʃi**k'**ala	'small'

- Fricatives: similar three-way contrast

		Initial		Medial	
Voiceless	s, ʃ, x, h	**x**ã	'scab'	i**x**a	'to laugh'
Voiced	z, ʒ, ɣ	**ɣ**ã	'messy hair'	ho**ɣ**ã	'fish'
Ejective/glottalized	s', ʃ', x'	**x'**ã	'to do'	ptu**x'**a	'to crumble'

- Sonorants: m, n, l, w, j

(6) Syllable structure

- Syllables generally of the form C_0V
- No codas morpheme-internally
  - Arise morpheme-finally through truncation, reduplication (also apocope in casual speech)
  - A couple exceptional affixes (*-l*, *-sh*)
- Extensive set of CC onset clusters
  - Unaspirated stop + sonorant: *blo*[1] 'potato', *gmũ* 'twisted'
  - Unaspirated stop + fricative: *kʃu* 'to bead', *ksapa* 'wise'
  - Fricative + C: *spã* 'to thaw', *xt͡ʃa* 'to bloom', *xmũ* 'to buzz', *ʃlo* 'to melt'
  - Nasal + nasal: *mni* 'water'
  - Stop + stop: *kte* 'to kill', *tka* 'heavy'

(7) Root shape

- Lakhota words can be quite long/morphologically complex

wa-	*ju-*	*o-*	*ki-*	*wãʒi-la*	'unite'
indef.	obj.-causation-	loc.-	to/for-one-	dimin.	

- Affixes tend to use simple syllable structure, unmarked segments (sonorants, stops)—e.g.,
  - Locative prefixes: *o-*, *i-*, *a-*, (*e-*), *ki-*
  - Valence changing prefixes: *wa-*, *a-*
  - Subject/object marking: *wa*, *ya*, *ũ(k)*, *ma*, *ni*, *tʃĩ*; reflexive ***-it͡ʃ'i-***
  - Types of causation: *ka-*, *na-*, *pa-*, *ya-*, *yu-*
  - Causative suffixes: *-ya*, *-kʰija*
  - Progressive: *-hã* (etc.)
- Clusters, fricatives, ejective, etc., are mostly found in roots, which are often quite short
  - C_0, or $C_0VC(V)$[2], C_0VC_0V
  - Longer monomorphemic roots: *paɣõta* 'duck', *wagmiza* 'corn', *it͡ʃamna* 'snow'

[1] Unaspirated stops become voiced before sonorants in clusters.

[2] There is a set of roots that are traditionally analyzed as /C_0VC/, but which always surface with an additional final vowel; see Shaw (1980) for discussion. The analysis of these 'C-final' roots is irrelevant to the restrictions discussed here.

  - However, long roots frequently behave as if they are morphologically complex (e.g., other morphemes are 'infixed' in the middle of them, or other words with shared pieces and vaguely related meanings; for discussion, see Boas and Deloria 1941, pp. 26–28)
- Root-affix asymmetry: Faith/ROOT ≫ *CC, *Laryngeal, *Fricative ≫ Faith
  - See Parker 1997, Beckman 1998, Alderete 2001
- This is important, because it already reduces the probability of encountering multiple clusters, ejectives, etc., in a word
  - Plausibly attributed to the morphological structure of words + phonological restrictions on affixes, not a ban on multiple clusters/fricatives/etc.

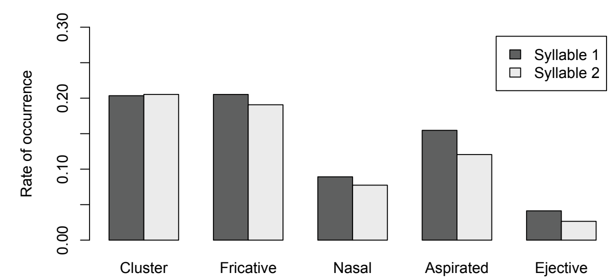
☞ Upshot: cooccurrence restrictions only interesting within roots, where marked structures are freely allowed

(8) Positional effects within roots: very weak

- Clusters, fricatives, aspirated stops/affricates, and ejectives can all occur in either initial or medial position

	Initial	Medial
Clusters	***sk**uja* 'sweet', ***gl**epa* 'vomit'	*hã**sk**a* 'tall', *hĩ**gl**a* 'move suddenly'
Fricatives	***ʃ**apa* 'dirty', ***ʒ**ata* 'forked'	*wa**ʃ**ĩ* 'work', *wã**ʒ**i* 'one'
Aspirates	***pʰ**ata* 'to butcher', ***kʰ**ãta* 'plum'	*a**pʰ**a* 'hit', *wa**kʰ**ã* 'holy'
Ejectives	***p'**o* 'fog'	*nã**p'**ĩ* 'wear around the neck'

- Nasals look rather infrequent when counted this way, but only 2 nasals, compared with 4 ejectives, 4 aspirates, 6 fricatives, and numerous clusters…
- Rough counts based on a database of 2,275 monosyllabic and disyllabic roots (details below)



- Rate of occurrence of clusters, fricatives, aspirated stops, nasals, and ejectives is not appreciably lower in medial position in disyllabic roots
- Thus, we expect that a reasonable number of roots should have these structures in both positions simultaneously

(9) Hints of cooccurrence restrictions

- (Richard T. Carter 1974, p. 51): cites claim by Robert Hollow (p.c.) of "one consonant cluster per morpheme"
  - Notes that there are exceptions—e.g., ***gleʃka*** 'spotted', ***wikcemna*** '10'
  - Nonetheless, appears to be at least a tendency
- Points out ejectives and aspirates pattern like clusters in this respect
  - Restriction on two aspirates, two ejectives, or combinations of the two is not surprising (MacEachern 1999)
  - However, implies that aspirates and ejectives also do not cooccur with clusters
- Informal observation suggests that roots with two fricatives are also scarce
  - Lots of words like *kaɣe* 'to make', *sapa* 'black', *ɣopa* 'snore'
  - Few words like *xu**ɣ**a* 'dent', *ʃoʃa* 'bubbly'
  - Virtually no words like **ɣoʃa* (distinct fricatives)
- Combinations of fricatives with clusters are also rare
  - A few cluster-first examples: ***mnuɣa*** 'eat crunchily', ***ʃloɣa*** 'make hominy'
  - Very few fricative-first examples: ***ɣoptã*** 'listen', [**ʃ**a**kp**e] 'six'
- Roots with ejective + cluster combinations difficult to find
  - ****k'omna***, ****tkuʃ'e***

(10) A preliminary check: Munro (1989) verb list

- Munro (1989): 1,075 verb entries
- Includes many morphologically complex entries, and many related forms
  - *o-mna* 'to smell (tr.)', *s'a-mna* 'to stink', *o-s'a-mna* 'to stink', *se(w)i-mna* 'to have a gamey smell', *ʃit͡ʃa-mna* 'to smell bad', *waʃte-mna* 'to smell good'
- Removed all prefixes (even if semantically opaque)
  - Also bound stems in obviously complex forms—e.g., *s'a-* in *s'a-mna* 'to stink' (no prefix or independent form *s'a*, but obviously contains *mna* 'smell')
- Result: 399 simple roots

(11) Observed vs. expected counts: fricative combinations

- Probability of disyllabic verbs: ≈57%
- Probability of fricatives: first syllable 32%, second syllable: 18%
- Expected two-fricative verb roots in this database: 23
- Actual number: 6
  - ***ʃloɣa*** 'to make hominy', ***xmuɣa*** 'to curse', ***ʃoʃe*** 'to froth', ***ʃkiʃka*** 'rough', *(pa-)**sisa*** 'to pin', *(ka-)**xuɣa*** 'to break'

- Gradient cooccurrence restriction on multiple fricatives
  - Not completely illegal, but underattested
  - Attested examples often involve complete/near identity: ʃkVʃkV, xVɣV

(12) Observed vs. expected counts: cluster combinations

- Probability of disyllabic verbs: ≈57%
- Probability of clusters: first syllable: 28%, second syllable: 12.5%
- Expected number of two-cluster verb roots in this database: 8
- Actual number: 4
  - ***gleʃka*** 'spotted', ***wiktʃemna*** 'ten', ***blaska*** 'flat', ***ʃkiʃka*** 'rough'

(13) Observed vs. expected counts: singleton fricative + cluster combinations

- Cluster-first: expected 21, vs. actual 6
  - ***gleɣa*** 'colorful', ***gleʃka*** 'spotted', ***gnaɣa*** 'burned up', ***mnuɣa*** 'eat crunchily', ***ʃkiʃka*** 'rough', ***ʃloɣa*** 'make hominy'
- Fricative-first: expected 9, vs. actual 2
  - ***ɣoptã*** 'listen', ***ʃakpe*** 'six'

☞ An unexpected interaction: fricatives tend not to cooccur with clusters??

(14) Summary so far

- Preliminary examination confirms prevalence of gradient cooccurrence restrictions
- OCP restrictions on structures that do not appear to give rise to categorical bans (?)
  - Fricatives: OCP($\begin{bmatrix} -\text{son} \\ +\text{cont} \end{bmatrix}$) ?
  - Clusters: OCP(CC)
- Restrictions on combinations of 'logically independent' structures
  - *CC… $\begin{bmatrix} -\text{son} \\ +\text{cont} \end{bmatrix}$, * $\begin{bmatrix} -\text{son} \\ +\text{cont} \end{bmatrix}$…CC
- Limitations
  - Smallish data set → low expected numbers of occurrences (hard to be sure that O/E of 4 out of 8 is meaningful)
  - Particularly limiting for aspirates, ejectives
- To address this issue, we turn to a larger data set

3 Probing these effects with a larger lexicon

(15) A larger lexicon: (Buechel and Manhart 2002)

- Constructed a database of 12,396 entries in Buechel and Manhart (2002) dictionary

- All entries except those beginning with *a*-, *i*-, *o*- (omitted since mostly prefixed, and could be left out without systematically undercounting any particular class of consonants)

- Rough morphological parsing to identify roots
  - Separated recurring roots and prefixes (distributional criterion)
  - Used meaning as check to reduce hyperanalysis
    * *hoɣã* 'fish' ≠ *ho*+*ɣã* 'bushy voice'; *tʰaspã* 'apple' ≠ *tʰa*+*spã* 'thaw ruminant's body part'
  - Much ambiguity remains concerning degree of synchronic decomposition
    * *tʰaxtʃa* 'deer' based on *tʰaɣa* 'spit/cud'??
    * *ũtʃiʃitʃala* 'crow' related to *ũtʃiʃi* 'mother-in-law'?
- Extracted unique morphemes (4,860)
  - Removed reduplicants and contracted forms (show independent phonological restrictions)
  - Removed allomorphs derived by final vowel changes ('ablaut')
  - Removed morphemes longer than 2 syllables, as more likely to be complex or misanalyzed
  - Result: 351 monosyllabic, 1924 disyllabic morphemes
- Caveat: many entries that are almost certainly morphologically complex remain undecomposed, and some words probably also over-decomposed
  - Less of a problem in 1–2 syllable morphemes, which we focus on here
  - Net effect would be to water down results concerning root-internal cooccurrences
- One other caveat: marking of aspiration in dictionary appears to be incomplete, especially for certain segments in initial position (e.g., [tʃ] vs. [tʃʰ])
  - I omit counts of aspirated stops in initial position

(16) Results: Observed/Expected counts

- As before, calculated rate of occurrence of consonant types in initial syllable (monosyllables and disyllables together) and second syllable (disyllables only)[3]
- Results show that some combinations occur about as often as one would expect based on the independent frequencies, while other combinations occur far less often

Initial	Medial							
	Cluster		Fricative		Aspirated		Ejective	
Nasal	27/35 =	✓0.77	39/33 =	✓1.19	20/21 =	✓0.97	8/5 =	✓1.76
Cluster	43/80	? 0.53	45/75	? 0.60	11/47	*? 0.23	1/10	*? 0.10
Fricative	43/81	? 0.53	54/75	? 0.72	13/48	*? 0.27	10/10 =	✓0.96
Ejective	5/16	*? 0.31	13/15 =	✓0.86	0/10	* 0.00	8/2 =	3.80

[3]This is the source of the data in the graph in (8).

(17) Making sense of this pattern

- Claim: degree of underattestation of a combination is related to the degree of grammatical dispreference for the individual elements involved
- Nasals are relatively unrestricted
  - Nasals are arguably marked relative to oral consonants (e.g., nasal place contrast doesn't occur without corresponding oral contrast; Maddieson 1980, p. 60)
  - However, in Lakhota, nasals can freely occur in roots and affixes, and are relatively frequent within roots (proportionally speaking)
  - No interaction: combinations with nasals about as frequent as expected (O/E > .75)
- Clusters and fricatives appear to be somewhat restricted
  - Typologically marked; occur only within roots in Lakhota
  - Also cooccur with themselves and with each other somewhat less often than expected (O/E between .5 and .75)
- Aspirates and ejectives are more restricted ⇒ stronger superadditive effect
  - Also occur (almost) exclusively within roots in Lakhota, and infrequently, at that
  - Clusters and fricatives cooccur with aspirates rather rarely (O/E ≈ .25)
  - Clusters and ejectives also rarely cooccur (O/E between .1 and .31)
  - Other interesting effects involving ejectives, not discussed here.[4]
- Strongest effect: ejectives and aspirates
  - May also be due to a more general ban on laryngeal cooccurrence (OCP(Lar.))

(18) Summarizing this result

- "Superadditive" effects: dispreferred elements cooccur even less often than expected, based on degree of dispreference for either one independently
- Such effects have been noted before in the literature
  - See especially Frisch (1996, p. 153) ('poor get poorer'); Albright (2008)
- Crucially, not as simple as saying that combinations of rare things are extra-rare
  - Nasals are relatively infrequent in Lakhota, but are not grammatically restricted; no superadditive effects
  - Restrictions on marked combinations go beyond joint probability of subparts

[4]The high rate of cooccurrence of ejectives with other ejectives turns out to be a by-product of a high rate of pseudoreduplicated morphemes involving ejectives (*s'ĩs'ĩ, t'ĩt'ĩ, x'ax'a, x'ãx'ã, x'ex'e*), some of which are mimetic. This may also partly explain the high rate of occurrence of ejectives with fricatives, since fricative ejectives were counted as fricatives in these counts, and a substantial proportion of the pseudoreduplicated roots involve fricatives. There does genuinely appear to be a preference for C'VSV roots, however: *k'eɣa* 'scrape', *k'oɣa* 'rattle', *k'eza* 'hard', *t'ĩza* 'stiff/tight', *t'oʒa* 'dull', etc. It is possible that this is related to another fact about the distribution of ejectives in Lakhota, namely, that they tend not to cooccur with stops of any sort (plain or ejective). The literature on laryngeal cooccurrence restrictions has generally focused on cases where cooccurring stops must either assimilate or dissimilate for ejection (MacEachern 1999). Gallagher (2009) hypothesizes that the reason for such patterns is that it is difficult to distinguish plain vs. ejective stops in the presence of an ejective; we speculate that this may favor fricatives (and sonorants, not counted here) in the presence of an ejective.

4 Proposal: ganging up of weighted constraints

(19) The challenge

- Allow some combinations of marked structures to occur as often as expected based on independent probability, while other combinations receive an additional penalty
- Amount of additional superadditive effect should follow from grammatical status of subparts
  - Combinations of things which are already marginal become even worse
- Effect is lexically gradient—perhaps universally so?
- A root/morpheme structure effect: no alternations in Lakhota (or any known case?)
- In this section, we generate a lexicon with these properties, using a grammar of weighted constraints in the style of Harmonic Grammar (Legendre, Miyata, and Smolensky 1990; Pater, in press)

(20) Step 1: the "trickle-down" base

- We need to get some combinations to occur *as often as expected*, and other combinations to occur *less often than expected*
- "Occur as often as expected" = baseline probability, given the frequency of segments involved
- Let's assume that potential words are generated by sampling segments in proportion to their relative frequency in the language
  - Flemming (2006) presents arguments that the input to phonological grammar is not the rich base directly, but rather the output of a set of contrast constraints that define the language-specific inventory
  - Here, we simply extend this to suppose that candidate words are generated by sampling from the inventory in a way that reflects their relative frequency[5] of the segments involved
- "Occur less often than expected" = some of these candidates will be penalized or eliminated by the phonological grammar, based on the combinations of segments that they contain

(21) Step 2: eliminating combinations with weighted constraints

- The goal: let constraints like *CC and *Fricative gang up to eliminate (or strongly penalize) candidates that combine clusters and fricatives
- We treat the elimination of possible words as losing to the null parse
  - Contemplated and rejected by Prince and Smolensky (1993/2004); see also Albright (2008)

[5]Another possibility is that the inventory grammar itself is probabilistic, and that frequency differences among segments drawn from the inventory are due to differences in phonetic markedness, rather than purely reflecting segment frequencies in the ambient language.

- E.g., schematically:
  - Fricatives allowed: Weight(MParse) > Weight(*Fricative)

/ɣã/ Weight:	MParse 4	*CC 3	*Fric 3	Total
☞ a. ɣã		-1		-3
b. *null*	-1			-4

  - Clusters allowed: Weight(MParse) > Weight(*CC)

/gla/ Weight:	MParse 4	*CC 3	*Fric 3	Total
☞ a. gla		-1		-3
b. *null*	-1			-4

  - Combinations not allowed: Weight(*CC) + Weight(*Fricative) > Weight(MParse)

/ɣãgla/ Weight:	MParse 4	*CC 3	*Fric 3	Total
a. ɣãgla		-1	-1	-6
☞ b. *null*	-1			-4

- Two issues
  - We need the effect to be gradient, not categorical
  - We need to ensure that the effect depends on the status of the individual elements involved

(22) Gradient well-formedness

- The tableaux above simply assume that the candidate with the smallest penalty wins
- Following Maximum Entropy approaches (Goldwater and Johnson 2003; Jaeger, to appear), we may instead assume the following:

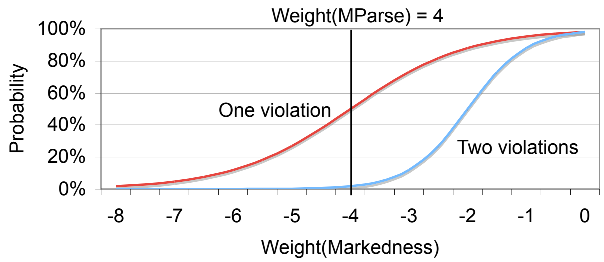
$$\text{Prob(output)} = \frac{e^{\text{penalty(output)}}}{\sum\limits_{\text{candidates}} e^{\text{penalty(candidate)}}}$$

  - Penalty = weighted sum of violations
- For the schematic examples above:
  - Prob(ɣã), Prob(gla) ≈ 73%
  - Prob(ɣãgla) ≈ 12%

(23) Step 3: penalized vs. non-penalized combinations

- Under the scheme in (22), the farther the weight of a markedness constraint is below MParse, the more likely the structure is to surface
  - Weight(MParse) ≫ Weight(Markedness)
- However, this condition also makes it relatively unlikely that the markedness constraint can gang up to overcome MParse (e.g., 2×Weight(Markedness) > Weight(MParse))

- Gang effects occur in a relatively small band around the weight of MParse



  - Discrepancy between probability of one vs. two violations is most noticeable when Weight(Markedness) is close enough to MParse that probability of single violation is also beginning to decline
- This captures the observation that superadditive effects are seen only for those markedness constraints that are strong enough to cause mild dispreference for structures independently
- Crucially, although frequency may be one factor governing the weight of markedness, other factors may be relevant, too
  - E.g., substantive biases (Wilson 2006)
  - Lack of superadditive effects involving *Nasal in Lakhota may reflect very low weighting of *Nasal (regardless of the fact that nasals are somewhat infrequent)

(24) No alternations motivated by cumulative markedness?

- The formulation here, in terms of MParse, reflects a grammar that is designed to answer the question 'what is the probability that word /X/ could surface'?
- One possibility: two modes of evaluation
- Phonotactic well-formedness: assess relative ranking of markedness w.r.t MParse, which sets a well-formedness threshhold
  - May gradually shape the lexicon over time (bad combinations are avoided; Martin 2007), but is not directly responsible for transforming ill-formed inputs into preferred outputs
- Deciding on a production output requires faithfulness, as in standard OT
  - One-for-one trade-off of markedness and faithfulness violations generally precludes alternations specifically to remedy cumulative complexity
- For a different approach, which allows cumulative effects with standard faithfulness constraints, see Albright, Magri, and Michaels (2008)

(25) Local summary

- Expected rates of occurrence derived by sampling inputs to the grammar according to the frequency of the segments involved

- Grammar (probabilistically) eliminates certain inputs, based on severity of markedness violations
  - Eliminates = selects the null parse instead (i.e., produces no overt output)
  - The fact that the null parse violates just one constraint (MParse) allows multiple markedness violations to gang up and produce cumulativity effects
- Strength of cumulative effect depends straightforwardly on strength of simpler effects
  - Summed violations of simpler constraints
  - Precludes large disparities between well-formedness of simple violations of multiple violations
  - Predicts primarily gradient effects

5 Conclusion

(26) Summary

- Analysis of cumulative complexity effects using weighted constraint
  - Somewhat more powerful model of constraint interaction, designed to capture unexpected interactions between apparently unrelated constraint violations
- Highly modular approach to contain/constraint predicted effects
  - Maximum entropy-style competition largely limits to gradient effects, targeting already marginal structures
  - Complexity effects limited to static phonotactics by deriving through interaction of MParse, rather than faithfulness
- These limitations attempt to capture what I take to be a significant observation that such effects do not surface as categorical effects driving alternations

(27) Some open issues

- Fuller investigation of interactions among constraints within Lakhota
  - Numbers reported here are a preliminary investigation using new database (under construction)
  - In principle, should make it possible to compare a wide variety of different types of structures, to observe which do or do not enter into cumulative effects
- Cross-linguistic prevalence of cumulative effects?
  - Complexity thresholds are often denied, but cumulative effects don't actually appear to be all that difficult to find (Frisch 1996; Albright 2008; various others)
  - If it's right that they're characteristically gradient, we may just not have had the tools needed to find and document them
- Relation between frequency and weighting of markedness contraints
  - Often taken to be rather direct relation (e.g., Zuraw 2000; Hayes and Wilson 2008)
  - Model proposed here permits certain degree of discrepancy between segment frequency and weighting of constraints on features (e.g., *Nasal)

– In this case, independent confirmation that this is plausible comes from the discrepancy between the lowish frequency of nasals vs. the fact that they occur freely even in affixes
– This is taken as motivation for deriving baseline frequencies through an independent mechanism (here, as part of inventory constraints)
– Further investigation is needed into division of labor in modeling frequency effects, and how and why constraint weights may become dissociated from the frequency of the structures that violate them

References

Albright, A. (2008). A universally gradient co-occurrence restriction? Talk presented at 16mfm.

Albright, A., G. Magri, and J. Michaels (2008). Modeling doubly marked lags with a split additive model. In H. Chan, H. Jacob, and E. Kapia (Eds.), *BUCLD 32: Proceedings of the 32nd Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

Alderete, J. D. (2001). *Morphologically Governed Accent in Optimality Theory*. New York: Routledge.

Beckman, J. (1998). *Positional Faithfulness*. Ph. D. thesis, UMass.

Boas, F. and E. Deloria (1941). *Dakota Grammar*, Volume 23 of *Memoirs of the National Academy of Sciences*. Washington: United States Government Printing Office.

Buechel, E. and P. Manhart (2002). *Lakota dictionary: Lakota-English/English-Lakota*. Lincoln, NE: University of Nebraska Press.

Coetzee, A. and J. Pater (2008). Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic. *Natural Language and Linguistic Theory 26*, 289–337.

Flemming, E. (2006). The role of distinctiveness constraints in phonology. MIT ms.

Frisch, S. (1996). *Similarity and Frequency in Phonology*. Ph. D. thesis, Northwestern University.

Frisch, S. A., J. B. Pierrehumbert, and M. B. Broe (2004). Similarity avoidance and the ocp. *Natural Language & Linguistic Theory 22*(1), 179–228.

Gallagher, G. (2009). The role of contrast in laryngeal cooccurrence restrictions. In *Proceedings of WCCFL 27*. Somerville, MA: Cascadilla Press.

Goldwater, S. and M. Johnson (2003). Learning OT constraint rankings using a maximum entropy model. In J. Spenader, A. Eriksson, and O. Dahl (Eds.), *Proceedings of the Workshop on Variation within Optimality Theory, Stockholm University*, pp. 111–120. Stockholm University.

Hayes, B. and C. Wilson (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry 39*, 379–440.

Jäger, G. (to appear). Maximum entropy models and Stochastic Optimality Theory. In J. Grimshaw, J. Maling, C. Manning, J. Simpson, and A. Zaenen (Eds.), *Architectures, Rules, and Preferences: A Festschrift for Joan Bresnan*. Stanford: CSLI Publications.

Kirchner, R. (1996). Synchronic chain shifts in Optimality Theory. *Linguistic Inquiry 27*, 341–351.

Legendre, G., Y. Miyata, and P. Smolensky (1990). Harmonic grammar: A formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations. Technical Report 90-5, Institute of Cognitive Science, Univ. of Colorado.

MacEachern, M. R. (1999). *Laryngeal Cooccurrence Restrictions*. Garland Publishing.

Maddieson, I. (1980). Phonological generalizations from the UCLA phonological segment inventory database. In *UPSID: The UCLA phonological segment inventory database*, Number 50 in UCLA Working Papers in Phonetics.

Martin, A. (2007). *The evolving lexicon*. Ph. D. thesis, UCLA.

Moreton, E. and P. Smolensky (2002). Typological consequences of local constraint conjunction. In L. Mikkelsen and C. Potts (Eds.), *WCCFL 21*, pp. 306–319. Cascadilla Press.

Munro, P. (1989). Lakhota verb list. Unpublished UCLA ms.

Parker, S. G. (1997). An OT account of laryngealization in Cuzco Quechua. In *Work Papers of the Summer Institute of Linguistics*, Volume 41, pp. 47–57. University of North Dakota.

Pater, J. (in press). Weighted constraints in generative linguistics. *Cognitive Science*.

Prince, A. and P. Smolensky (1997). Optimality: From neural networks to universal grammar. *Science 275*, 1604–1610.

Prince, A. and P. Smolensky (2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell Publishing.

Richard T. Carter, J. (1974). *Teton Dakota Phonology*. Ph. D. thesis, University of New Mexico.

Rose, S. and R. Walker (2004). A typology of consonant agreement as correspondence. *Language 80*, 475–531. http://www-rcf.usc.edu/~rwalker/ABC.pdf.

Shaw, P. A. (1980). *Theoretical Issues in Dakota Phonology and Morphology*. New York: Garland Publishing, Inc.

Walter, M. A. (2007). *Repetition Avoidance in Human Language*. Ph. D. thesis, MIT.

Wilson, C. (2006). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science: A Multidisciplinary Journal 30*, 945–982.

Zuraw, K. (2000). *Patterned Exceptions in Phonology*. Ph. D. thesis, UCLA.