LSA–Boston
Jan. 10, 2004

Adam Albright
albright@ucsc.edu

The Emergence of the Marked: Root-Domain Markedness in Lakhota

1 Introduction

(1) The basic observation: marked structures are distributed unevenly throughout language, with "strong" or "privileged" positions frequently allowing a greater range of structures

- *Positional neutralization* (Trubetzkoy 1939 (1962); Steriade 1994; Jun 1995; Padgett 1995; Steriade 1997; Casali 1997; Beckman 1998; Lombardi 1999; Zhang 2002; Barnes 2002; Alderete 2003; Steriade, in press; and many others)

Onsets Stressed syls Roots Root-initial syls	tend to allow a wider range of structures than	codas unstressed syls affixes or reduplicants root-final syls

(2) The standard recipe for handling this asymmetry in OT:

1. Provide constraints that refer specifically to strong/privileged positions, such as:
   - ➢ Positional faithfulness (Casali 1997; Beckman 1998): IDENT$_{Root}$(constricted glottis) (preserve ejectives in roots), etc.
   - ➢ Positional markedness (Zoll 1998; Smith 2002)
2. High-ranked positional constraints determine the range of structures in strong positions

/t'apa/	IDENT$_{Root}$(c.g.)	*[+c.g.]
☞ a. t'apa		*
b. tapa	*!	

3. The markedness constraint *[+c.g.] is outranked, but still present in the grammar; it can still play a role when IDENT$_{Root}$(c.g.) is not relevant (such as outside roots)

   E.g., hypothetical suffix /-t'a/:

/pala-t'a/	IDENT$_{Root}$(c.g.)	*[+c.g.]	IDENT(c.g.)
a. pala-t'a		*!	
☞ b. pala-ta			*

- ➢ The emergence of the unmarked (TETU; McCarthy and Prince 1994)): lower ranked markedness constraints get their way when special faithfulness constraints are inapplicable

(3) Asymmetries in possible structures are a natural consequence of two assumptions OT:

- Universal constraint set contains constraints that are designed to license/preserve a greater range of contrasts in privileged positions (see esp. Smith 2003)
- Outranked constraints are nonetheless active in the grammar, waiting to have an effect with higher-ranked positional constraints are inapplicable

(4) Perhaps for this reason, such cases have received a good deal of attention in the OT literature

- Cuzco Quechua: plain, aspirated, and ejective stops in roots, but only plain stops in suffixes (Parker and Weber 1996; Beckman 1998)
- Nootka: codas allowed in general, but prohibited in reduplicants (discussed by McCarthy and Prince 1994))

(5) Purpose of this paper: call attention to a less-discussed pattern, in which strong positions have systematically *simpler* structures than weak positions

(6) Lakhota: codas are generally banned in roots, but they may surface in affixes, function words, and reduplicants (data below)

➢ Proposal: such patterns are best captured by markedness constraints whose domain in the root (RSCs)

(7) Outline of the paper

- Description of Lakhota: possible roots, affixes, clitics, and reduplicants
- The same set of codas is allowed everywhere . . . except roots
- Root-domain Structure Constraints: markedness constraints whose domain of application is the root (or lexical category)
- Discussion: the motivation of root-specific constraints on syllable structure, and some unresolved issues

2 The distribution of codas in Lakhota

(8) Preliminaries on Lakhota

- Siouan language, spoken primarily in N. and S. Dakota, Nebraska, Minnesota, and Canada
- Data from Boas and Deloria (1941), Buechel (1970), Shaw (1980), Munro (1989), and field work with a native speaker, Mary Rose Iron Teeth
- Consonant inventory

unaspirated	p	t	tʃ	k
aspirated	pʰ[1]	tʰ	tʃʰ	kʰ
ejective	p’	t’	tʃ’	k’
fricatives		s, z, s’	ʃ, ʒ, ʃ’	x, ɣ, x’
nasals	m	n		ŋ
liquid		l		
glides			j	w

(9) Lakhota words allow a fairly rich set of onsets (A representative selection)

Stop + stop:

pte	‘cow’
tke	‘heavy’
tkʰa	‘but’
ktʰũ	‘wear’

Stop + fric/affric:

psĩ	‘rice’
pʃa	‘sneeze’
kʃto	(emph. clitic)
kt͡ʃi	‘with’

Fric + stop:

xtætu	‘evening’
xpæ	‘lie down’
stu	‘in love’
ʃkate	‘play’

Obstruent + sonorant:

blo[2]	‘potato’
gli	‘arrive home’
gnæ	‘cheat, fool’
sni	‘cold’

Nasal + nasal:

mni	‘water’

(10) Word-final codas are generally *not* permitted[3]

- (*[kat], *[tax], *[man])

(11) Word-medially: arguably also no codas

- Same clusters allowed medially as initially
- VCCV syllabified V.CCV in deliberate/slow repetition: [ja. . . tke]

[1] Aspiration is frequently realized with velar, rather than laryngel frication, particularly before [o], [a], and [ã], but frequently also before [e] and [u].

[2] UR most likely /plo/; voiceless unaspirated stops become voiced before sonorants.

[3] Here and throughout, I mean that codas *are not permitted on the surface.* Traditional analyses of Lakhota do make use of URs with underlying codas (CVC roots); see Shaw (1980) for arguments. Nothing that follows here will depend crucially on whether or not there are underlyingly C-final roots in Lakhota.

(12) There are a a handful of words with codas, however
(Again, not utterly complete, but this seems to be a large percentage of them)

[l]	ˈel	'in, at, to'
	maˈhel	'on'
	ˈlel, ˈhel	'here', 'there'
	tuˈktel	'somewhere'
	tuˈktektel	'here and there'
	eˈt͡ʃʰel, ˈhet͡ʃʰel	'that way, thus'
	tʰãˈkal	'outside'
	ˈkʰul	'under, beneath, down' (cf: kʰuta 'low down')
[n][4]	eˈhan, heˈhan	'at that time', 'then'
	leˈhan	'now, at this time'
	toˈhan	'when, until'
	iˈt͡ʃʰan, xt͡ʃʰeˈhan	'just then'
	ˈxta.le.han	'yesterday'
[m]	i.ˈsam, ˈsam	'more'
[s]	he.ˈnɔs	'they two'
[ʃ]	eniʃ, ˈniʃ, naʔĩʃ	'or'
	miʃ, niʃ, iʃ, ũˈkiʃ	'I', 'you', 'he/she/it', 'we' (emphatic/contrastive)
	ˈhet͡ʃeʃ	'that's how (it happened) (cf: ˈhet͡ʃel 'thus')
	ˈataʃ	'now'
	naˈkũʃ	'also'
	aˈkʰeʃ	'but, although, though'
	kaˈkʰeʃ	'at any rate'
	ˈk'ejaʃ	'but, although; some'
	waˈnakaʃ	'long ago'
	tkʰaʃ	'but' (dubitative) (cf: tkʰa 'but')
[x]	hũx	'some'
[p][5]	iˈtʰokap	'before' (time and place)
[k]	tak	'what' (cf: also ˈtaku)
	ĩʃˈtok	'is it?' (tag question) (cf: also ĩʃˈtoka)

(13) Two observations about these words:

- They are all function words of some sort (and there may be some relation between the place words in [l], and also the time words in [n]; [ʃ] seems to be emphatic/contrastive)
- The codas they contain are relatively 'good codas' (mostly acoustic continuants, absolutely no voiced/aspirated/glottalized obstruents)

(14) Furthermore: codas in suffixed/cliticized words in casual speech

- Animate plural *-pi*

Apocope to coda [p]:	/juha-pi/	'have-PL'	→	[juhap]
	/t͡ʃʰãˈzeka-pi/	'angry-PL'	→	[t͡ʃʰãˈzeka-p]
Or [m] after nasal V:	/lowã-pi/	'sing-PL'	→	[lowãm]
	/jatkã-pi/	'drink-PL'	→	[jatkãm]
Or [w] before nasals:	/oki-pi na/	'can-PL and'	→	[okiwna]

[4] These are written in Lakhota orthography as <l> after nasal vowels (*ehanl*); I do not know whether there is a more conservative pronunciation with [l].

[5] Phonetically, coda [p] (and also [k]) often have a small and variable amount of voicing, particularly in certain contexts, such as before [h], and utterance-finally. I assume this is a purely phonetic effect, and continue to transcribe them as voiceless [p]/[k]. It should be noted that Lakhota generally lacks surface [b],[g] except in stop+resonant clusters.

- Definite determiner *-ki*

  Apocope to coda [k]: /juˈha-ki/ ‘have-DEF’ → [juˈhak]
  /ˈjute ki/ ‘eat DEF’ → [ˈjutek]

- Emphatic clitic *-kʃto* → *-kʃ*

  /niˈje/ ‘you’ + /-kʃ/ → [nijekʃ] ‘it’s up to you’
  /ˈle͡ʃija na-wa-ʒĩ/ ‘here stand-1sg’ + /-kʃ/ → [ˈle͡ʃija ˈnɔʒĩkʃ] ‘I’m standing over here!’

- Other clitics don’t undergo apocope (low vowels, or bad codas):

-hã	progressive
-kta/-kte	future
-tkʰa	dubitative
-ʃni	negative
-t͡ʃa	definite
-he	addressing 2nd person

- Also happens to be no [-ti] suffix to create coda [t]

(15) Consistent set of “marginally possible codas”: [p,k,s,ʃ,x,m,n,l,w]

- Systematically absent: voiced, aspirated, and glottalized obstruents
  (*[z, ʒ, ɣ, pʰ, tʰ, kʰ, p’, t’, k’, s’, ʃ’, x’])
- Curiously also absent: [t] (perhaps an accidental gap because no *-ti* suffix?)

(16) What to make of this set of codas?

- Could ignore the problem
  - ➢ A relatively small handful of exceptions, after all, compared to the otherwise robust generalization that codas are disallowed in Lakhota
- Could try to contain the problem
  - ➢ Admit that they are exceptional, and find a way to regulate and contain the exceptionality
  - ➢ E.g., greater faithfulness for high frequency items, protected from the regular grammar by their frequency of occurrence?
    - ◇ *Laryngeal coda $\gg \mathcal{F}$ if frequent enough $\gg$ *Coda
    - ◇ See Zuraw (2000) for a proposal for handling exceptions, and why high frequency might help them resist the ordinary grammar; see also Bybee (2001) for discussion from a different perspective
  - ➢ Rests on the (almost certainly false) hope that all of these functional elements are more frequent than any lexical root
- Could analyze the problem as a perversion
  - ➢ Codas in suffixes and function words are preserved by special faithfulness constraints ($\mathcal{F}_{\text{Affix}}$, $\mathcal{F}_{\text{Func}}$)
  - ➢ For some reason, in Lakhota, $\mathcal{F}$ to these usually weak positions outranks $\mathcal{F}$ to roots[6]
- Or, turn the problem around
  - ➢ Propose that Lakhota *does* generally allow codas
  - ➢ But banned in roots by a special version *CODA that applies only in roots (*$\text{CODA}_{\text{ROOT}}$)

Strong evidence for this last approach from reduplication

[6] A related, but distinct approach which I will not explore here is an analysis that employs cophonologies (Orgun 1996; Inkelas 1998; Inkelas and Zoll 2003): one for roots, and one for function words and affixes. The reduplication data in the following section would be especially problematic for a cophonology approach, since it would require the non-root phonology to apply to reduplicants, while root phonology applies to their bases.

(17) Reduplication in Lakhota: two types

- Final reduplication: [gle.ʃka] → [gle.ʃka.ʃka] 'spotted'
- Penult reduplication: [gle.ɣe] → [gle.gle.ɣe] 'colorful' (*[gle.ɣe.ɣe])

➢ The analysis of this distinction in OT difficult and not directly relevant here. See Shaw (1980), Marantz (1982), and Sietsema (1988) for rule-based approaches, and Nelson (2003) for an attempt to recast this analysis in OT; see also Albright (2002, chap. 5) and Hogoboom (2003) for discussion.

(18) Penult reduplication: frequently copies the following onset, yielding $[C_1VC_2]_\sigma$ reduplicant

[sa.pe]	→	[sa**p**.sa.pe]	'black'
[ʃa.pe]	→	[ʃa**p**.ʃa.pe]	'dirty'
[sa.ke]	→	[sa**k**.sa.ke]	'hard'
[t͡ʃʰã.ze.ke]	→	[t͡ʃʰã.ze**k**.ze.ke]	'angry'
[wa.ʃ'a.ke]	→	[wa.ʃ'a**k**.ʃ'a.ke]	'strong'

(19) When can reduplicants have closed syllables?

1. When C_2 is among the set of "marginally possible codas" (as in [sa**p**.sa.pe], [sa**k**.sa.ke])
2. When C_2 can be devoiced, deglottalized, or depalatalized to create a marginally possible coda

[s]	[pu.ze]	→	[pu**s**.pu.ze]	'dry'
	[ble.ze]	→	[ble**s**.ble.ze]	'clear'
[ʃ]	[pa.tʰu.ʒe]	→	[pa.tʰu**ʃ**.tʰu.ʒ.e]	'bend over'
	[ka.p'o.ʒe.la]	→	[ka.p'o**ʃ**.p'o.ʒe.la]	
[x]	[ka.ɣe]	→	[ka**x**.ka.ɣe]	'do, make'
	[pi.ɣe]	→	[pi**x**.pi.ɣe]	'boil'
[k]	[t͡ʃi.k'a.la]	→	[t͡ʃi**k**.t͡ʃi.k'a.la][7]	'small'
	[ʃi.t͡ʃe]	→	[ʃi**k**.ʃi.t͡ʃe]	'bad'

3. And remarkably, [t] → [l]

[l]	[kʰa.te]	→	[kʰa**l**.kʰa.te]	'hot'
	[ʃka.te]	→	[ʃka**l**.ʃka.te]	'play'
	[ʔo.ta]	→	[ʔo**l**.ʔo.ta]	'be many'

Or, occasionally [t] → [k]

[l]	[su.ta]	→	[su**k**.su.ta]	'hard'

4. Even stranger: sometimes irregular [t], [l] → [k]

[k]	[o.ʒu.la]	→	[o.ʒu**k**.ʒu.la]	'full'
	[a.ju.ta]	→	[a.ju**k**.ju.ta]	'look at'

➢ The mappings employed are not 100% regular (a phenomenon reminiscent of Turkish emphatic reduplication; Kelepir (2000, Wedel (2000)). Nonetheless, C_2 must always be a member of the set [p,k,s,ʃ,x,m,n,l,w] if it is to be copied.

(20) Further details: some variability

- C_2 is (almost) never copied when a medial CCC would result[8]

*[psk]	[ka.ska.pe]	→	[ka.ska.ska.pe]	(*[ka.skap.ska.pe])	'slap'
*[sgm]	[yu.gmu.ze]	→	[yu.gmu.gmu.ze]	(*[yu.gmu.gmu.ze])	'twist'
*[kxw]	[ka.xwo.ke]	→	[ka.xwo.xwo.ke]	(*[ka.xwok.xwo.ke])	'blow'

- Occasionally, C_2 resists copying even when it could be transformed into a possible coda

[xo.te] → [xo.xo.te] (*[xol.xo.te]) 'gray'

[7] The *-la* is a diminutive suffix that usually does not count when determining which syllable to reduplicate. (That is, [t͡ʃik.t͡ʃi.k'a.la] is at some level penultimate reduplication, not a third pattern of initial or antepenultimate reduplication.)

[8] I have recorded one exception: *bleze* → *blesbleze* 'clear'

(21) What is important for present purposes:

- Penultimate reduplication does often create closed-syllable reduplicants, and they actively employ the same set of codas seen in suffixes and function words

(22) How do we know that C_2 of the C_1VC_2 reduplicant is really a coda?

- For what it's worth, slow, deliberate speech: [sak] . . . [sa] . . . [ke]
- In many cases the resulting CC cluster is not a possible onset
  - ➢ [kap'oʃp'oʒela] could not to be syllabified [o.ʃp'o]
  - ➢ Similarly [pat^{h}uʃt^{h}uʒe] (*ʃt^h), [t͡ʃ^{h}ãzekzeke] (*kz), [k^{h}alk^{h}ate] (*lk^h), etc.
- Phonetic realization: reduplicated C_2 [p] and [k] often realized partly or fully voiced, as with coda [p] and [k] in word/utterance-final position
- Logic: why would C_2 be transformed to a possible coda, if it was not a coda?

(23) Upshot: Lakhota exhibits "The emergence of the marked"

- A set of codas that are allowed everywhere *except* roots: [p,k,s,ʃ,x,m,n,l,w]
- Affixes, function words, and reduplicants allow systematically *more* marked structures than roots do

3 Analysis using root-specific markedness

(24) How to capture emergence of the marked?

- Possibility 1: greater faithfulness in weak positions (emergence of the unmarked in roots)
- Possibility 2: greater markedness in strong positions (emergence of the marked outside roots)

Goal of this section: argue against a faithfulness-based approach, and in favor of markedness relativized to roots

(25) Sketch of a faithfulness-based approach:

- Inviolable phonotactics: no laryngeally specified codas (*[z, ʒ, ɣ, p^h, t^h, k^h, p', t', k', s', ʃ', x']), and also no [t] (*t]$_\sigma$?)
- $\mathcal{F}$ for weak positions, to preserve codas in these positions, as long as they are a member of the OK set
- *CODA: ranked low enough that codas are allowed in weak positions
- $\mathcal{F}$ for other positions: ranked even lower, codas banned in roots

➢ Emergence of the marked arises when $\mathcal{F}$ for weak positions is irrelevant, and *CODA prevails

(26) Problems with the faithfulness-based approach:

- Relies on ranking $\mathcal{F}$ for weak positions $\gg$ $\mathcal{F}$ for strong positions (argued by some to be impossible, either because $\mathcal{F}_{\text{Weak}}$ does not exist, or a universal ranking of $\mathcal{F}_{\text{Strong}} \gg \mathcal{F}_{\text{Weak}}$; Casali 1997; Beckman 1998; Alderete 2001; Alderete 2003)
- Also: what constraint enforces faithfulness for weak positions?
  - ➢ Affixes: $\mathcal{F}_{\text{Aff}}$ (or general $\mathcal{F}$)
    Function words: $\mathcal{F}_{\text{Func}}$
    Reduplicants: MAX$_{\text{BR}}$
  - ➢ The only way to enforce identical phonotactics in all three positions is to ensure that *all three* of these faithfulness constraints are ranked identically w.r.t. coda markedness (below ban on laryngeal codas, above *CODA). This is the OT equivalent of a conspiracy: three constraints must push in the same direction to achieve a unified surface pattern.

(27) Markedness-based approach mitigates both of these problems

- A limited set of codas are generally allowed in the language, using the familiar configuration of layered markedness and faithfulness (as above, I use broad cover terms like *Laryngeal Coda to stand in for your favorite implementation of coda conditions)

  ➢ *Laryngeal Coda ≫ IO-$\mathcal{F}$, MAX_{BR} ≫ *CODA

  Roots like /sap/ cannot surface faithfully:

/mahel/		*Laryngeal Coda	IO-$\mathcal{F}$	MAX_{BR}	*CODA
☞	a. [mahel]				*
	b. [mahel*V*]/[mahe][9]		*!		

  Affixes may also contain codas:

/ʃa-p/		*Laryngeal Coda	IO-$\mathcal{F}$	MAX_{BR}	*CODA
☞	a. [ʃap]				*
	b. [ʃap*V*]/[ʃa]		*!		

  As may reduplicants:

/RED-sape/		*Laryngeal Coda	IO-$\mathcal{F}$	MAX_{BR}	*CODA
☞	a. [sapsape]			*	*
	b. [sasape]			**!	

  But hypothetical function word /maheʒ/ cannot surface as such (choice of output depends on details of IO-$\mathcal{F}$

/maheʒ/		*Laryngeal Coda	IO-$\mathcal{F}$	MAX_{BR}	*CODA
	a. [maheʒ]	*!			*
☞	b. [maheʃ]/[maheʒe]/[mahe]		*		(*)

- And codas are banned altogether in roots

  ➢ $\text{*CODA}_{\text{ROOT}}$ ≫ IO-$\mathcal{F}$

  Function words like /tuktel/ surface faithfully:

/sap/		$\text{*CODA}_{\text{ROOT}}$	*Laryngeal Coda	IO-$\mathcal{F}$	MAX_{BR}	*CODA
	a. [sap]	*!			*	
☞	b. [sape]/[sa]			*		

(28) Mirror image of the TETU configuration: here, greater faithfulness emerges outside roots, when a higher ranked markedness constraint ($\text{*CODA}_{\text{ROOT}}$) is inapplicable

[9] As is often the case with richness of the base problems, data from the language do not necessarily tell us what happens to non-occurring patterns (i.e., what happens to clicks in English?). In fact, the standard analysis of Lakhota claims that illegal codas are fixed by epenthesis (Shaw 1980), so I will limit the candidates to those with the most obvious fixes of epenthesis or deletion.

4 Discussion and conclusion

(29) This is, in fact, not the first proposal that strong positions may be subject to extra markedness

- Pre-OT: Morpheme Structure Rules/Conditions could be stated on any position of any morpheme (Halle 1959; Stanley 1967; Chomsky and Halle 1968)
- de Lacy (2000), Smith (2002): strong positions are sometimes required to "sound strong"
  - ➢ Stressed syllables must be heavy (weight-to-stress)
  - ➢ Vowels lengthen in stressed syllables (*Short V/ˊσ)
  - ➢ Word-initial onsets must be voiceless (extremely non-sonorant)

(30) But if strong positions can be subject to extra markedness, don't we lose the predicted asymmetry ((3) above)?

➢ Potentially yes. But ...

- One line of defense: limit what kinds of markedness can apply in strong positions
  - ➢ de Lacy, Smith: only those motivated by sonority considerations[10]
  - ➢ The Lakhota example shows that things other than sonority may play a role as well; but we still might hope that it's not the entire set of $\mathcal{M}$
- In order to answer this question (why *CODA$_{\text{ROOT}}$, and what else might we need?), it would be helpful to have a better understanding of what motivates such cases

(31) A suggestion about why syllable structure, from the history of Lakhota

- A speculative, but plausible origin of this RSC in Lakhota: (Rood 1983)

  (Modern Lakhota forms stand in for the equivalent Proto-Siouan reconstructed forms)

  Stage 1: roots may end in C or V, clitics are V-initial (w/deletion of clitic V in hiatus)

C-final roots:	/sak/	sak-eʔ	sak-api	sak-ĩkte
V-final roots:	/gleʃka/	gleʃka-ʔ	gleʃka-pi	gleʃka-kte
	/waʃte/	waʃte-ʔ	waʃte-pi	waʃte-kte

  Stage 2: C-final roots reanalyzed as variant V-final roots

"C-final roots":	sak{a∼e∼ĩ}	sake-ʔ	saka-pi	sakĩ-kte
V-final roots:	/gleʃka/	gleʃka-ʔ	gleʃka-pi	gleʃka-kte

- RSCs on syllable structure are caused by (or serve to facilitate) morphological parsing

(32) More generally: constraints on syllable structure and sequences play an important role in parsing

- Regularities in possible sequences aid in learning word/morpheme segmentation (Harrington, Watson, and Cooper 1988; Cairns, Shillcock, Chater, and Levy 1994; Brent and Cartwright 1996; Aslin, Woodward, LaMendola, and Bever 1996; Saffran, Newport, and Aslin 1996; Cairns, Shillcock, Chater, and Levy 1997; Brent 1999; Aslin, Saffran, and Newport 1999)
- Hay (2003) and others have argued that these cues are not only useful in learning, but are actively used by the adult parser as well

(33) A possible parallel in English: root-final obstruent clusters

	Root-finally	Word-finally
Voiceless-final clusters generally OK	✓st	✓st
	✓kt	✓kt
Voiced-final clusters bad in roots	*gd[11]	✓gd
	*zd	✓zd

- (The full data is, of course, more complex than this)

[10] Inkelas and Rose (2003) argue that positional velar fronting in child speech, in which velar stops are realized as alveolar in strong positions, is also an example of strengthening gone awry: the child attempts to produce a strengthened [k], but because of articulatory constraints of having a big tongue and a small palate, the result of over-vigorous ends up being closer to [t].

[11] A possible exception is *smaragd* 'emerald'

(34) RSCs employing inventory constraints as well

- Possible initial segments in Welsh roots (shaded) vs. words: (Williams 1980; Thorne 1993)

b	d	g	p	t	k			
m	n	ŋ	m̥ʰ	n̥ʰ	m̥ʰ			
v	ð		f	θ	s	x	x^w	h
	l, r			ɬ, $r̥^h$				

(35) Before we can really decide powerful a theory of root markedness is needed, we need a better empirical understanding of root structure constraints that cannot be reduced to surface phonetic constraints

- Also relevant to consider to what extent these constraints parallel/are different from possible restrictions on particular subsets of roots (lexical strata: Itô and Mester 1995; Inkelas, Orgun, and Zoll 1997; Itô and Mester 1999)

(36) Where this all leads:

- It is not a completely unrestricted call to allow any sort of positional markedness constraint in the grammar
- RSCs like the Lakhota case have plausible *morphological grounding*, in the same way that many constraints have been argued to have *phonetic grounding* (Hayes 1999; Smith 2003)

(37) An open question:

- Can all strong positions be referred to by such markedness constraints?
  - ➢ Roots (Lakhota)
  - ➢ Initial onsets (Welsh)
  - ➢ Stressed syllables? onsets in general? Seem unlikely... (and more difficult to motivate on parsing grounds)

  Conjecture: should yield predominantly *edge* phenomena (see Broselow 2003 for some relevant cases)

(38) One other important issue that I leave unresolved here: the learning issue posed by RSCs

- A subset problem: the set of structures found in Lakhota roots are a subset of those found in the language as a whole (open syls $\subset$ all syls)
- A learner, hearing the word [$tk^haʃ$], cannot respond by demoting assuming that codas are allowed *everywhere*
  - ➢ There may be no positive evidence that forces the learner to go back and conclude that codas are impossible in roots
  - ➢ Result: an inadequately restrictive grammar
- This problem may be complicated even further by the fact that learners do not necessarily have a complete morphological analysis (what is a root and what is not) when they are beginning to learn such phonotactic distributions
- Work underway trying to determine to what extent this can already be handled by existing proposals, such as Biased Constraint Demotion (BCD; Prince and Tesar 1999)

(39) Conclusion

- An overlooked, but perhaps quite common pattern: roots have systematically simpler structures than words as a whole (the emergence of the marked outside roots)
- Lakhota: roots may not have codas, but affixes, function words, and reduplicants may
- Proposal: a wider range of markedness constraints may target strong positions than previously assumed
  - ➢ Not a negative result! Not only is it empirically necessary, but it opens an avenue to explore the ways in which markedness is employed to help in tasks like morphological parsing, beyond the obvious sequential constraints that have been explored thus far

References

Albright, A. (2002). *The Identification of Bases in Morphological Paradigms*. Ph. D. thesis, UCLA.

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

Alderete, J. D. (2003). Structural disparities in Navajo word domains: A case for LEXCAT-FAITHFULNESS. *The Linguistic Review 20*, 111157.

Aslin, R., J. Woodward, N. LaMendola, and T. Bever (1996). Models of word segmentation in fluent maternal speech to infants. In J. Morgan and K. Demuth (Eds.), *Signal to Syntax: Bootstrapping From Speech to Grammar in Early Acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Aslin, R. N., J. R. Saffran, and E. L. Newport (1999). Statistical learning in linguistic and nonlinguistic domains. In B. MacWhinney (Ed.), *Emergence of Language*. Hillsdale, NJ: Lawrence Earlbaum Associates.

Barnes, J. (2002). *Positional Neutralization: A Phonologization Approach to Typological Patterns*. Ph. D. thesis, UC Berkeley.

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.

Brent, M. (1999). Speech segmentation and word discovery: A computational perspective. *Trends in Cognitive Science 3*, 294–301.

Brent, M. and T. Cartwright (1996). Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition 61*, 93–125.

Broselow, E. (2003). Marginal phonology: Phonotactics on the edge. *The Linguistic Review 20*, 159193.

Buechel, E. (1970). *A Dictionary of Teton Sioux*. Pine Ridge, S.D.: Red Cloud Indian School, Inc.

Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

Cairns, P., R. Shillcock, N. Chater, and J. Levy (1994). Lexical segmentation: the role of sequential statistics in supervised and unsupervised models. In *Proceedings of the 16th Annual Meeting of the Cognitive Science Society [August 1994]*, pp. 136–141.

Cairns, P., R. Shillcock, N. Chater, and J. Levy (1997). Bootstrapping word boundaries: A bottom-up corpus-based approach to speech segmentation. *Cognitive Psychology 33*, 111–153.

Casali, R. F. (1997). Vowel elision in hiatus contexts: Which vowel goes? *Language 73*, 493–533.

Chomsky, N. and M. Halle (1968). *The sound pattern of English*. New York: Harper and Row.

de Lacy, P. (2000). Markedness in prominent positions. In *Paper presented at the HUMIT (Harvard-MIT Student Conference in Linguistics)*. http://www.cus.cam.ac.uk/~pvd22/Docs/handouts/delacy-humit-talk.pdf.

Halle, M. (1959). *The Sound Pattern of Russian*. The Hague: Mouton.

Harrington, J., G. Watson, and M. Cooper (1988). Word boundary identification from phoneme sequence constraints in automatic continuous speech recognition. In *Proceedings of the 12th International Conference on Computational Linguistics*, Hillsdale, NJ, pp. 225–230. Erlbaum.

Hay, J. (2003). *Causes and Consequences of Word Structure*. New York: Routledge.

Hayes, B. (1999). Phonetically-driven phonology: The role of Optimality Theory and inductive grounding. In M. Darnell, E. Moravscik, M. Noonan, F. Newmeyer, and K. Wheatly (Eds.), *Functionalism and Formalism in Linguistics, Volume I: General Papers*, pp. 243–285. Amsterdam: John Benjamins.

Hogoboom, A. (2003). Reduplicant placement, anchoring and locality. UCSC ms.

Inkelas, S. (1998). The theoretical status of morphologically conditioned phonology: a case study from dominance. *Yearbook of Morphology 1997*, 121–155.

Inkelas, S., O. Orgun, and C. Zoll (1997). The implications of lexical exceptions for the nature of grammar. In I. Roca (Ed.), *Derivations and Constraints in Phonology*, pp. 393–418. Oxford University Press.

Inkelas, S. and Y. Rose (2003). Velar fronting revisited. In B. Beachley, A. Brown, and F. Conlin (Eds.), *BUCLD 27: Proceedings of the 27th annual Boston University Conference on Language Development*, pp. 334–345. Somerville, MA: Cascadilla Press.

Inkelas, S. and C. Zoll (2003). Is grammar dependence real? UC Berkeley and MIT ms. (Rutgers Optimality Archive #587).

Itô, J. and A. Mester (1995). The core-periphery structure of the lexicon and constraints on reranking. In J. Beckman, S. Urbanczyk, and L. Walsh (Eds.), *University of Massachusetts Occasional Papers in Linguistics [UMOP] 18: Papers in Optimality Theory*, pp. 181–209. Amherst: GLSA.

Itô, J. and A. Mester (1999). The structure of the phonological lexicon. In N. Tsujimura (Ed.), *The Handbook of Japanese Linguistics*, pp. 62–100. Oxford: Blackwell Publishers.

Jun, J. (1995). *Perceptual and articulatory factors in place assimilation: an Optimality Theoretic approach:*. Ph. D. thesis, UCLA.

Kelepir, M. (2000). To be or not to be faithful. In A. Göksel and C. Kerslake (Eds.), *Studies on Turkish and Turkic Languages*. Harrassowitz Verlag.

Lombardi, L. (1999). Positional faithfulness and voicing assimilation in Optimality Theory. *Natural Language and Linguistic Theory 17*, 267–302.

Marantz, A. (1982). Re Reduplication. *Linguistic Inquiry 13*(3), 435–482.

McCarthy, J. and A. Prince (1994). The emergence of the unmarked: Optimality in prosodic morphology. In M. Gonzàlez (Ed.), *Proceedings of the North East Linguistics Society 24*, Amherst, MA, pp. 333–379. GLSA.

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

Nelson, N. (2003). *Asymmetric Anchoring*. Ph. D. thesis, Rutgers.

Orgun, O. (1996). *Sign-based morphology and phonology: with special attention to Optimality Theory*. Ph. D. thesis, UC Berkeley.

Padgett, J. (1995). Partial class behavior and nasal place assimilation. In *Proceedings of the South Western Optimality Theory Workshop 1995*, pp. 145–183. Tucson, AZ: The University of Arizona Coyote Papers.

Parker, S. G. and D. Weber (1996). Glottalized and aspirated stops in Cuzco Quechua. *International Journal of American Linguistics 62*, 70–85.

Prince, A. and B. Tesar (1999). Learning phonotactic distributions (technical report ruccs-tr-54). Technical report, Rutgers Center for Cognitive Science, Rutgers University.

Rood, D. S. (1983). A preliminary consideration of Proto-Siouan ablaut. *Na'pao: A Saskatchewan Anthropology Journal (Special Issue: Proceedings of the Second Siouan Languages Conference, 1982) 13*, 24–28.

Saffran, J. R., E. L. Newport, and R. N. Aslin (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language 35*, 606–621.

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

Sietsema, B. (1988). Reduplications in dakota. In G. L. L. McCleod and D. Brentari (Eds.), *CLS 24: Papers from the 24th Annual Regional Meeting of the Chicago Linguistic Society. Part One: The General Session*, Chicago. The Chicago Linguistics Society.

Smith, J. (2002). *Phonological Augmentation in Prominent Positions*. Ph. D. thesis, UMass.

Smith, J. (2003). Towards a compositional treatment of positional constraints: The case of positional augmentation. In A. Carpenter, A. Coetzee, and P. de Lacy (Eds.), *UMOP26: Papers in Optimality Theory II*, pp. 337–370. Amherst, MA: GLSA.

Stanley, R. (1967). Redundancy rules in phonology. *Language 43*, 393–436.

Steriade, D. (1997). Phonetics in phonology: The case of laryngeal neutralization. UCLA ms.

Steriade, D. (in press). Directional asymmetries in place assimilation: a perceptual account. In E. Hume and K. Johnson (Eds.), *Perception in Phonology*. Academic Press.

Thorne, D. (1993). *A Comprehensive Welsh Grammar/Gramadeg Cymraeg Cynhwysfawr*. Cambridge, MA: Blackwell.

Trubetzkoy, N. S. (1962). *Grundzüge der Phonologie, 3rd ed.* Göttingen: Vandenhoeck & Ruprecht.

Wedel, A. (2000). Perceptual distinctiveness in Turkish emphatic reduplication. In R. Billerey and B. D. Lillehaugen (Eds.), *WCCFL 19: Proceedings of the 19th West Coast Conference on Formal Linguistics*, Somerville, MA, pp. 546–559. Cascadilla Press.

Williams, S. J. (1980). *A Welsh Grammar*. Cardiff: University of Wales Press.

Zhang, J. (2002). *The Effects of Duration & Sonority on Contour Tone Distribution: A Typological Survey and Formal Analysis*. New York: Routledge.

Zoll, C. (1998). Positional asymmetries and licensing. MIT ms.

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