24.941J / 6.543J / 9.587J / HST.727J The Lexicon and Its Features
Spring 2007

Summary from last time

- Expected findings:
  - reliance on a feature set shapes rule learning/generalization
  - evidence of generalization by interpolation, not extrapolation
- What we did find:
  - extrapolation does occur: HH extends to back and lax V's
  - when it doesn't it's maybe for independent reasons: FH does not extend to vowels at the other end of the height continuum from the targets.
  - We expected at least *feature-dependent* extrapolation, but the reported results don't speak to this.
- Significance:
  - Unclear what the evidence says about the role of features in learning rules/constraints.

An experiment to test the role of features in rule learning

- Make the rule decidedly unnatural: e.g. eliminate any connection between context and change.

  (They can be learned: Dell et al. 2000 JExPs and others)
- Use as a context class a large set of segments and its subsets: e.g. continuant C's (± son, ±strident, ±voiced, ±anterior)
- Vary conditions only as a function of the nature of the sets of conditioning segments.
- Use several natural and many unnatural conditions.

Compare these rules: contexts describable in F-terms

- l -> n/ {θ, ð, s, ʃ, z, ʒ, f, v} V _
- l -> n/ {θ, ð, f, v} V _
- l -> n/ {s, ʃ, z, ʒ} V _
- l -> n/ {θ, s, ʃ, f} V _
- l -> n/ {θ, ð, s, z} V _
- l -> n/ {ð, z, ʒ, v} V _
- l -> n/ {ð, v} V _

With these rules

- l -> n/ {θ, s, ʃ, z, f, v} V _ [- ð, ʒ]
- l -> n/ {θ, ð, v} V _ [- f]
- l -> n/ {s, ʃ, ʒ} V _ [-z]
- l -> n/ {s, f} V _ [- ʃ, θ]
- l -> n/ {θ, z} V _ [- ð, s, ʃ, ʒ]

Two results to look for

- Learning differences as a function of the feature-based analysis of the context.
- Extrapolation patterns:

  if this rule is generalized

  l -> n/ {ð, v} V _

  is it generalized randomly or only to natural classes of context C's?

Natural classes

‘sets like [p, t, k], [m, n, ɳ, ŋ], and [i, e, ä] are classes of sounds that appear together in rule after rule; sets like [e, x, n], [ʔ, r, f, w], or [a, s, b] are rarely (if ever) found in any rules in any language’ (p. 22). The widely attested sets are referred to as natural classes. It has been observed that natural classes can generally be given simple phonetic characterizations, e.g. [p, t, k] are voiceless stops, [m, n, ɳ, ŋ] are nasal stops, and [i, e, ä] are front vowels.

Kenstowicz and Kisseberth (1977), apud Flemming 2004.

The lore on natural classes

- The grammar regulates the occurrence of feature values, not of segments.
- Sets of sounds are targetted together in a process only because they share the feature value targetted.
- Same for context.
- But the feature set is not known a priori. To discover it, we study the natural classes in attested systems.
- (there are other ways: Lindblom, Stevens).

Issues

- Universality of phonological features and categories:

  do all grammars function with a single set of features?
- Deviations from universality:

  are features induced from observed patterns of sound distribution?
- What are the right methods to address these issues?

Two empirical studies of natural classes

- Mielke (2004 OSU diss., Phonology 2005, WCCFL 2005):
  n languages, 561 processes.
- Hinskens and van der Weijer (Lx 2003): distribution of consonantal modifications in the 315 languages of Maddieson 1984
- The same question: how common are the processes that refer to unnatural classes.

Hinskens and van de Weijer

- Checking the natural class status of sets of C's subject to consonantal modifications (MOD) in Maddieson 1984 (315 languages)
- MOD types:
  - aspiration, ejection, breathy voicing
  - pre/post nasalization
  - palatalization, velarization, labialization, pharyngealization

Coding scheme

- The feature set in Clements and Hume 1995
- A natural class cannot have one member only
- The description of a natural class can employ only conjunctions of existing feature values
- If the F is privative, the non-F class is not natural.
- The natural class cannot be the complement of a class described by some conjunction of F-values.
- Mild counterexamples: the class of C's subject to some MOD contains a gap of just one C not subject to MOD
- Serious counterexamples: all others.

Tolowa vs. Wintu

T	**p**	**t**	**k**	**kw**	
	±	**t'**	**k'**	**kw'**	
W	**p**	**t**	**k**		**q**
	p^h	**t^h**	**-**		**-**

Hypothesis 1: MOD occur on sets forming natural classes

	LAR						ORA				NAS		
Hyp 1	asp	brr	brv	eje	lar	pra	lab	pal	pha	vel	pm	noMOD	Total
+	76	2	6	31	13	2	30	5	1	2	14	0	182
±	10	0	0	15	2	0	8	6	1	1	2	0	45
–	7	0	1	5	11	0	8	9	4	4	2	0	51
noMODhl	3	0	0	0	0	0	0	0	0	0	0	150	153
Total	96	2	7	51	26	2	46	20	6	7	18	150	431

The relationship between the 11 types of MOD (as well as the absence of MOD) and the evaluation of Hypothesis 1

Figure by MIT OpenCourseWare.

Analysis

- H&vdW use 3 types of constraints:
  - X/MOD: Class X C's have MOD.
  - *MOD: no segment has MOD
  - Ident MOD IO: preserve input MOD

(a) X/MOD >> *MOD, Ident MOD:

all C's in X have MOD: no contrast in class C

(b) Ident MOD >> X/MOD >> *MOD

UR MOD contrasts are preserved: contrast in class C

(c) *MOD >> X/MOD, Ident MOD

No segments have MOD; no contrast.

- In this analysis, the mild and serious counterexamples have to be dealt with using unspecified exception mechanisms.
- The one-vs.-more-gap coding scheme looks irrational.
- The ranking in (a) seems ok for some MOD types (aspiration) but overgenerates for most others (pharyngealization, labialization, ejection).
- Anything is a possible X/MOD constraint if X is a natural class: this also overgenerates (e.g. [+strident]/prenasalized; [+anterior]/breathy voiced)
- The real focus here: a look at the cases labeled as counterexamples suggests that nearly all are highly systematic. We just need more constraints.

- If the system is rational, then only some C features will interfere with MODs.
- Specifically, only C-features whose cues a given MOD type interacts with will give rise to apparent exceptions from natural-class behavior of that MOD type.
- So palatalization/velarization might be expected to interact with primary place
- But prenasalization or length are not expected to interact with primary place.

Non-interaction: prenasalization vs. primary place

- 18 languages with prenasalized C's
- 5 apparent counterexamples to hypothesis of non-interaction
  - Paez: stops and sibilant fricatives are ^{N}C, not affricates
  - Apinaye: all plosives can be ^{N}C, but not tʃ.
    There is a ᶮɟ lacking an oral counterpart
  - Mazateco: all voiced stops but [b] (loan) are ^{N}C.
  - Kewa: unique labial plosive is **mb**, there is a **d/nd** contrast, and no other plosive (**c**) is ^{N}C.
  - Nambakaengo: see next slide
- Only the Kewa and Nambakaengo gaps lack obvious explanation

Nambakaengo C's

p	**pw**	**pj**	**t**	**tw**	**tj**	**k**	**kw**	**kj**
p^h			**t^h**			**k^h**		
mb	**mbw**		**nd**	**ndw**		**ŋg**	**ŋgw**	
v			**s**					
m	**mw**		**n**	**nw**	**nj**	**ŋ**	**ŋw**	
								j

Non-interaction: length vs. primary place

- 11 languages with long C's (vastly underestimated)
- No counterexamples to the hypothesis of non-interaction
- Sundry odd restrictions
  - Shilha: pharyngealized can't be long (?)
  - X: Affricates can't be long (maybe they're too long already)
  - Iraqw: only s is distinctively long (?)
  - Maranunku: ɲ, ŋ, tʃ can't be long (they arise in clusters only?)
  - Wichita: s, ts contrast for length but not ts', tsh or any other C (?)

Post-aspiration

- This is a MOD type that is expected to interact with the place features of the C.
- We're testing a hypothesis that's stronger than H&vdW's natural class hypothesis:
  - all plosives or none contrast for aspiration
  - deviations from this can only be due to factors that directly affect the realization of the aspiration contrast.

Generalizations about distribution of aspiration contrasts in plosives

- 84 languages with C/C^h contrast.
- 67 (79%) have the contrast for all stops, regardless of point of articulation.
- 70 languages with C/C^h contrast and with affricates.
- 50 (71%) have the contrast for all plosives, regardless of affrication.

Generalizations about distribution of aspiration contrasts in plosives

- 79% of the languages have the C/C^h contrast equally available at all points of articulation
- The only clear deviation from the norm is lawful: C's further back may lack the contrast.
- 71% of the languages have the C/C^h contrast equally available for all plosives.
- The only clear deviation from this norm is also lawful: affricates may lack the contrast.
- Lack of contrast is variably realized as C or C^h.

	p	t̪	t	c	k	k^w	kp	q
Somali Navajo								
Viet. Tolowa								
Acoma								
Wintu								
Tiddim Chin								
Ewe,								
Hindi								
Beembe Zuni								

blue: C's that can't contrast for C^h; others do.

green: C's that must be aspirated; others can't be

light blue: expected to pattern like blue/green but don't occur

Nambakaengo C's again

p	**pw**	**pj**	**t**	**tw**	**tj**	**k**	**kw**	**kj**
p^h			**t^h**			**k^h**		
mb	**mbw**		**nd**	**ndw**		**ŋg**	**ŋgw**	
v			**s**					
m	**mw**		**n**	**nw**	**nj**	**ŋ**	**ŋw**	
								j

A lawful pattern

- Languages lacking p/p^h may have just a marginal p.
- Setting these aside:

  a. If the primary constriction of C_1 is behind that of C_2, then, if an h-contrast is possible for C_2, it is also possible for C_1.

  b. If C^j/C^w contrasts for h, its plain counterpart also does.

  c. A C that does not contrast for aspiration can be realized (or recorded) as either C or as C^h

- Reasons:

  (a) VOT increases as the oral constriction goes towards the back. For C's with constrictions further back, C^h would have to have a substantially longer VOT to contrast with its C counterpart. (b) C/ _ Glide has longer VOT than / _V.

Engstrand, Krull and Lindblom 2000, *PERILUS*

Nonsense syllables C_iVC_i in different stress conditions; Swedish

Subject	Place	Voiced focal		Voiceless focal		Voiceless contrastive		Voiceless emphatic	
		Mean	Standard deviations	Mean	Standard deviations	Mean	Standard deviations	Mean	Standard deviations
PT	LAB	9	9	52	12	72	14	100	25
	DEN	23	6	69	13	90	13	105	16
	VEL	29	11	77	14	102	19	105	19
JS	LAB	7	5	70	9	98	16	128	31
	DEN	15	4	71	16	158	39	223	42
	VEL	23	6	80	11	154	30	225	60
RL	LAB	1	2	69	9	89	15	233	52
	DEN	16	5	75	15	81	13	290	66
	VEL	27	28	89	13	97	12	351	45

VOT means and standard deviations (ms) for labial, dental and velar stops produced under the four conditions.

Blue: the C does occur and can't be aspirated.
Striped blue: if the affricate occurs it can't be aspirated; other C's are
Striped green: if the affricate occurs, it must be aspirated.
Blank C's: other plosives in the language contrast for aspiration

	pf	t̪	tθ	ts	tʃ	cç	tl	kx	Other C
Telugu									
Yakut-Kabardian									
Haida, Navajo, Hupa									
Zuni, Amuesha, Gilyak, Nama									

This is probably lawful too

- Affricates pattern exactly like the back constrictions: if they deviate from the general C/C^h contrast, they are reported as unaspirated or as invariably and non-contrastively aspirated.
- There are no languages where affricates only have the C/C^h contrast.
- C/C^h contrast on affricates may be harder for the same reason why it's impossible on fricatives: segmenting VOT from the frication at release.
- Laminals tend to be affricated, even if less noticeably.

Conclusion on H&vdW

- Deviations from natural class behavior:

  for the cases I studied, there are fewer than they report, maybe none.
- The right constraints distributing MOD's:

  these are are not MOD/X, for arbitrary (but natural) X classes, but constraints on the goodness of MOD contrast. More on this later.
- Preference for maximal use of MOD:

  A priori, one wouldn't expect just 21% of the languages that have a C/C^h contrast to restrict C's with naturally long VOT's from using it. If one constraint favors the contrast and one goes against it in some class of C's, one would expect an even distribution of aspiration-restricted and unrestricted languages. So there is maybe an additional factor that favors the across-the-board use of MOD contrasts: cf. Clements 2005.

Transition to Mielke

- There are lawful gaps in the classes of C's that express a specific MOD.
- This suggests that the surface inventory of C's arises from the interaction of constraints promoting use of MOD (Flemming's 1995 MAX Contrast) and specific constraints prohibiting a MOD contrast in some subclass (Flemming's 1995 MIN Dist).
- The key point now: the surface inventory does not provide an unobstructed view into the grammar: one has to unravel the interactions to recover the natural classes.
- Next: what conclusions about natural classes and features follow when this point is not heeded.

E. Sapir

- distributional similarity as the possible basis for a phonological category

Photograph removed due to copyright restrictions.

Nootka

ph	**th**	**tsh**	**tʃh**	**kh**	**kwh**	**qh**
p	**t**	**ts**	**tʃ**	**k**	**kw**	**q**
m	**n**	**l**	**j**		**w**	
p’	**t’**	**ts’**	**tʃ’**	**k’**	**kw’**	**q’**
’m	**’n**	**’l**	**’j**		**’w**	

Sapir 1933, reprinted in Mandelbaum ed. 1963

Phonetic differences

- **p! is an ejective**: “synchronous closure of lips and glottal cords […] sudden release of lip closure, a moment of pause and then the release of glottal closure […] click-like character”
- **’m is a preglottalized sonorant**: “while lip closure and glottal closure are synchronous as before, the glottal closure must be released at the point of initial sonancy of the **m**.”
- Spelling “difference **p!** vs. **’m** […] was not unjustified on purely phonetic grounds”

Alex Thomas

- Taught <**p!ap!i**:> but <**’ma:’mi:qsu**>
- Accepts <**p!ap!i:**>
- Volunteers <**m!a:m!i:qsu**>
- “valuable evidence for the phonologic reality of a glottalized class of consonants, which included both type **p’** (with prior release of oral closure) and type **’m** (with prior release of glottal closure).”
- basis for choosing this broader class, when a narrower one was suggested by the spelling?

Sapir: distributional parallelism

- Neither T’ nor ’R can occur syllable finally.
- Suffixes that turn T into T’ and R into ’R:
- **wi:nap** ‘stay’
- **wi:nap’-a/a** ‘stay on the rocks’
- **tlum** ‘to be hot’
- **tlu’m-a/a** ‘be hot on the rocks’

Language specific sound classes

- "Morphology… supports the phonologic proportion **p:p'= m:'m**[…] In other languages, with different phonologic and morphologic understandings, such a parallel of orthography might not be justified at all and the phonetic differences that actually obtain between **'m** and **p'** would have a significantly different psychologic weighting"

Yokuts

- Both T' and 'R can occur syllable finally.
- So can all other C's: no natural class here
- Suffixes turn R into 'R, but not T into T'
- '-feature seeks R, ignoring intervening T

(8) *Glottalisation of sonorants over another consonant*

ʔilk-	ʔel'kaahin	'sing'
dull-	dol'laahin	'climb'
yawl-	yaw̓laahin	'follow'

From Howe and Pulleyblank 2001, *Phonology*

- Suppose T’ and ’R share a phonetic feature: [+constricted glottis]
- Evidence was consistent with 2 classes:

  [-son, +c.g] and [+son, +c.g.]
- Spelling encouraged Alex to focus on 2 classes: <p!> vs. <’m>
- [spelling effect on analysis in <melon>, <cello>]
- Alex disregards spelling, glottal timing difference, probable auditory difference between T’ and ’R, to focus just on a shared articulatory property.

Why

- because features refer primarily to constrictions in the vocal tract - not timing or auditory properties? (cf. Browman and Goldstein)
- or because Alex is enforcing feature economy? (cf. Clements: reducing the feature/segment ratio)
- or underspecification? (cf. Clements, Lahiri)
- or because sounds are categorized primarily on the evidence of *distributional* similarities?

  (Sapir, and now Mielke and others).

Mielke

- Evaluates the success rate of 3 theories of DF’s (occasionally augmented) in characterizing natural classes in a survey of the rules found in 561 languages
- best theory (SPE) characterizes as featurally unified categories only 67% of the sets of triggers and targets found in these rules.
- some sets cannot be captured by any feature set or combinations of feature sets

Theory	*Preliminaries to Speech Analysis* (Jakobson, Fant, and Halle 1954)	*The Sound Pattern of English* (Chomsky and Halle 1968)	*Unified Feature Theory* (Clements 1990, Hume 1994, Clements and Hume 1995)
Features	***10 binary acoustically-defined features:*** (1) vocalic/non-vocalic (2) consonantal/non-consonantal (3) interrupted/continuant (4) checked/unchecked (5) strident/mellow (6) voiced/unvoiced (7) flat/sharp (8) grave/acute (9) tense/lax (10) nasal/oral ***1 equipollent acoustically-defined feature:*** (11) compact/diffuse	***23 binary articulatorily-defined features:*** (1) consonantal (2) vocalic (3) sonorant (4) continuant (5) voice (6) nasal (7) coronal (8) anterior (9) strident (10) lateral (11) back (12) low (13) high (14) round (15) distributed (16) covered (17) syllabic (18) tense (19) delayed primary release (20) delayed release of secondary closure (21) glottal (tertiary) closure (22) heightened subglottal pressure (23) movement of glottal closure	***17 binary features (effectively):*** (1) sonorant (2) approximant (3) vocoid (4) nasal (5) ATR (6) strident (7) spread (8) constricted (9) voice (10) continuant (11) lateral (12-14) anterior (C-place/V-place/either) (15-17) distributed (C-place/V-place/either) ***18 unary features:*** (18) C-place (19) vocalic (20) V-place (21-23) pharyngeal (C-place/V-place/either) (24-26) labial (C-place/V-place/either) (27-29) lingual (C-place/V-place/either) (30-32) dorsal (C-place/V-place/either) (33-35) coronal (C-place/V-place/either) ***Potentially unlimited binary aperture features:*** (36) open l, [(37) open2...]

Figure by MIT OpenCourseWare.

Pero stop assimilation

[-voice] or [+nasal] C triggers gemination

p	**t**			**k**	
b	d			g	
ɓ	ɗ				
		tʃ			
		dʒ			
m	**n**			**ŋ**	
	r				
	l				
			j		w

Figure by MIT OpenCourseWare.

Gemination vs. epenthesis

/káp/ + /kò/	→	[kákkò]	'he told'
/pét/ + /nà/	→	[pénnà]	'he went out'
/tʃúp/ + /kò/	→	[tʃókkò]	'he has shown'
/tʃɪrép/ + /mù/	→	[tʃírémmù]	'our women'

/káp/ + /dʒí/	→	[kávídʒì]	'eat (habitual)'
/tʃúg/ + /dʒí/	→	[tʃúgídʒí]	'talk (habitual)'

Figure by MIT OpenCourseWare.

The rule Mielke has in mind, after Frajzyngier

- Gemination [-syllabic] {[+nasal], [-voice]}

  [-syllabic] → Ø

  {[+nasal], [-voice]} → [+long]

If this is the right rule, what does its existence mean?

- Mielke: feature categories (e.g. stops, nasals, fricatives) arise in individual grammars from the observation of shared phonological patterning.
- They may or may not have a phonetic basis.
- When they do, that's because phonetically based sound changes (e.g. mergers caused by perceptual confusion) have created certain distributional restrictions.
- When they don't, that's because accumulated sound changes have obscured the phonetic factors behind original sound changes and the resulting distributions.

Emergent Feature Theory

- Learner acquires rules like Evenki nasalization in the same way as Japanese devoicing or rendaku:
- Identifies a shared structural change
- Enumerates observed target and context segments
- Members of such a list share a common diacritic, related or not to some shared phonetic property.
- {**p, t, k, m, n**} set defines a Pero-specific “feature”, learned by Evenki speakers from evidence of gemination.

The list is phonetically arbitrary

- “Labeling the property is a task primarily for linguists. Since these segments share a clear phonetic property, linguists may refer to this abstract property as something like [flat], [+round] or [Labial] […] to reflect phonetic similarity. However, since the phonetic similarity is secondary to the fact that the grouping is phonologically significant, the class could just as easily be thought of as“the segments that do X” and the abstract property that connects them could just as easily be called “z”. Phonological features and phonologically active classes are potentially isomorphic. No feature needs to be learned that is not motivated by the presence of a phonologically active class.” (p.116)

- Alternatively, here are some possible sources for unnatural class behavior:
  - Interactions with Faithfulness
  - Multiple, related Markedness constraints (Flemming 2004)
  - We lack the right features
  - We lack the right level of description (e.g. broad transcription is taken to represent a phonetic analysis of the sounds)

Pero stop assimilation

[-voice] or [+nasal] C triggers gemination

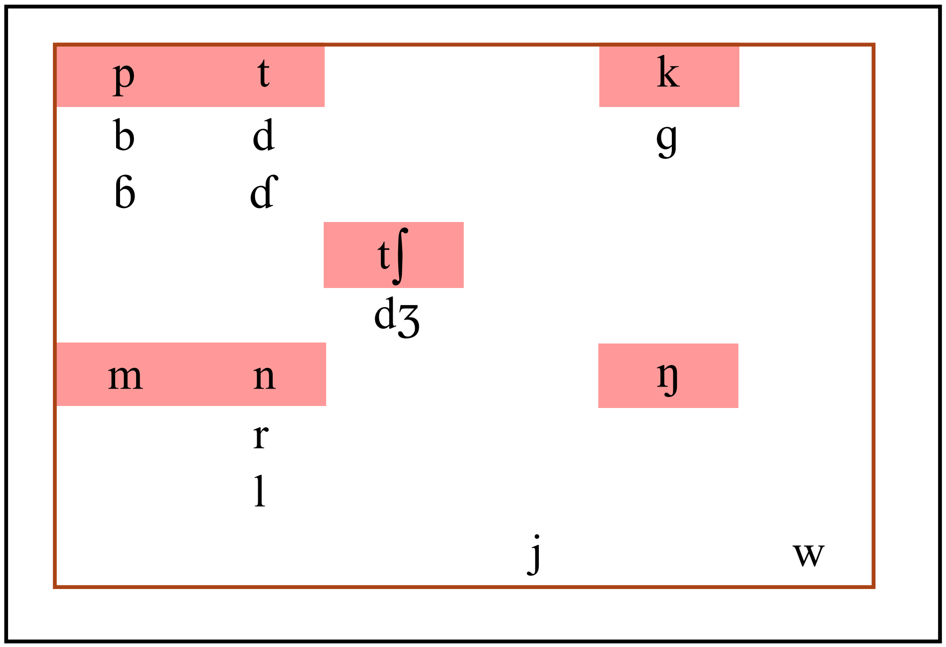


Figure by MIT OpenCourseWare.

An alternative for Pero

- V-insertion Ø -> i/C_ [+voice, -son]

- Gemination [-syllabic] [-syllabic]

  [-syllabic] → Ø

  [-syllabic] → [+long]

- "In V_V, the contrast between T and D is neutralized by the voicing of voiceless stops, in some environments, and by their spirantization in others." (F. 1989:17)

  tʃek-i [tʃigi] 'be lost', pet-i [piɾi] 'go out'
- Epenthetic V, like underlying V, triggers voicing/spirant.

  kap-dʒi [kavidʒi] 'eat-habitual'
- Some Epenthetic V's induce voicing then delete

  piit-ko 'make fire': **piit-u-ko** -> **piiɾugo** -> [**piiɾgo]**

  kavidʒi might come from **/kap-tʃi/,** w/t V-deletion **?**
- Now compare

  loop-ni [luubini], *[**luunni]**, 'beat him'

  tʃirep-mu [**tʃiremmu**], *[**tʃirebimu**]

CC clusters that can undergo gemination =
(total CC clusters) - (CC undergoing epenthesis).

UR	loop-ni	tʃirep-mu	kap-tʃi	kap-ko
epenthesis	loopini	n/a	kapitʃi	n/a
voicing	loobini	n/a	kavidʒi	n/a
gemination	n/a	tʃiremmu	n/a	kakko

- Given how common rule interaction is, it’s surprising then that only 20-30% (or less) of the sets of triggers and undergoer segments cannot be given a featural characterization **in the grammar’s total output.**

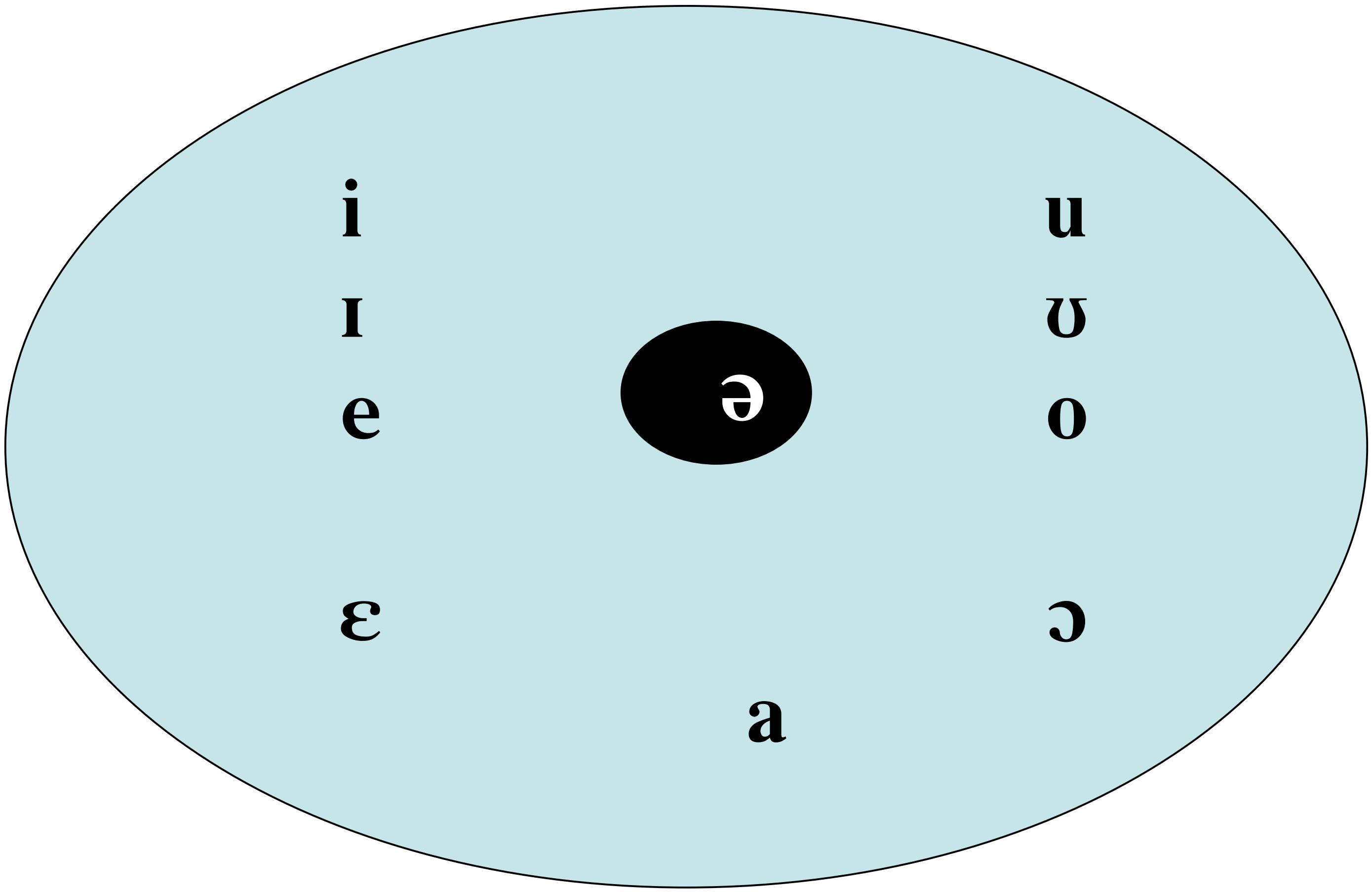
  One would expect a larger number.

  This is similar to the unexpectedly large number of inventories where MOD applies to the largest applicable class.

Next case

Natural classes emerge when we factor into the analysis the effect of segment-specific constraints and segment- or feature-specific *faithfulness* constraints.

Natural class obtained by excluding [ə]



stress

- Can be carried by {a, e, i….u} but not [ə]
- French: last V ≠ [ə] is stressed
- Dutch, English, Indonesian: unstressable
- Property setting [ə] apart: short, not expandable

	diph	long	i,u,y	short	ə
open	140	128	81	78	57
closed	120	107	84	68	50

Dutch: Klabbers&Van Santen, ICSLP 2000

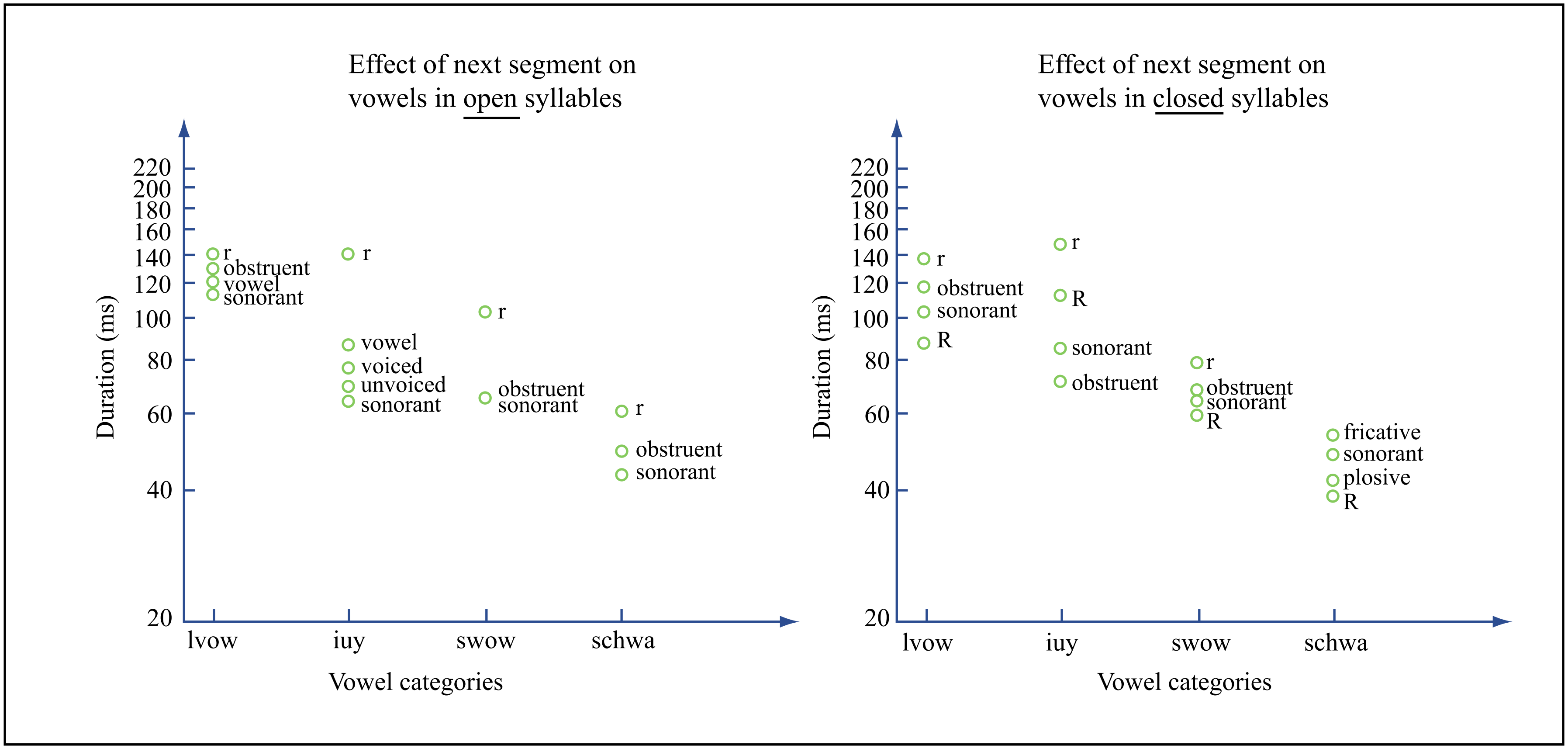


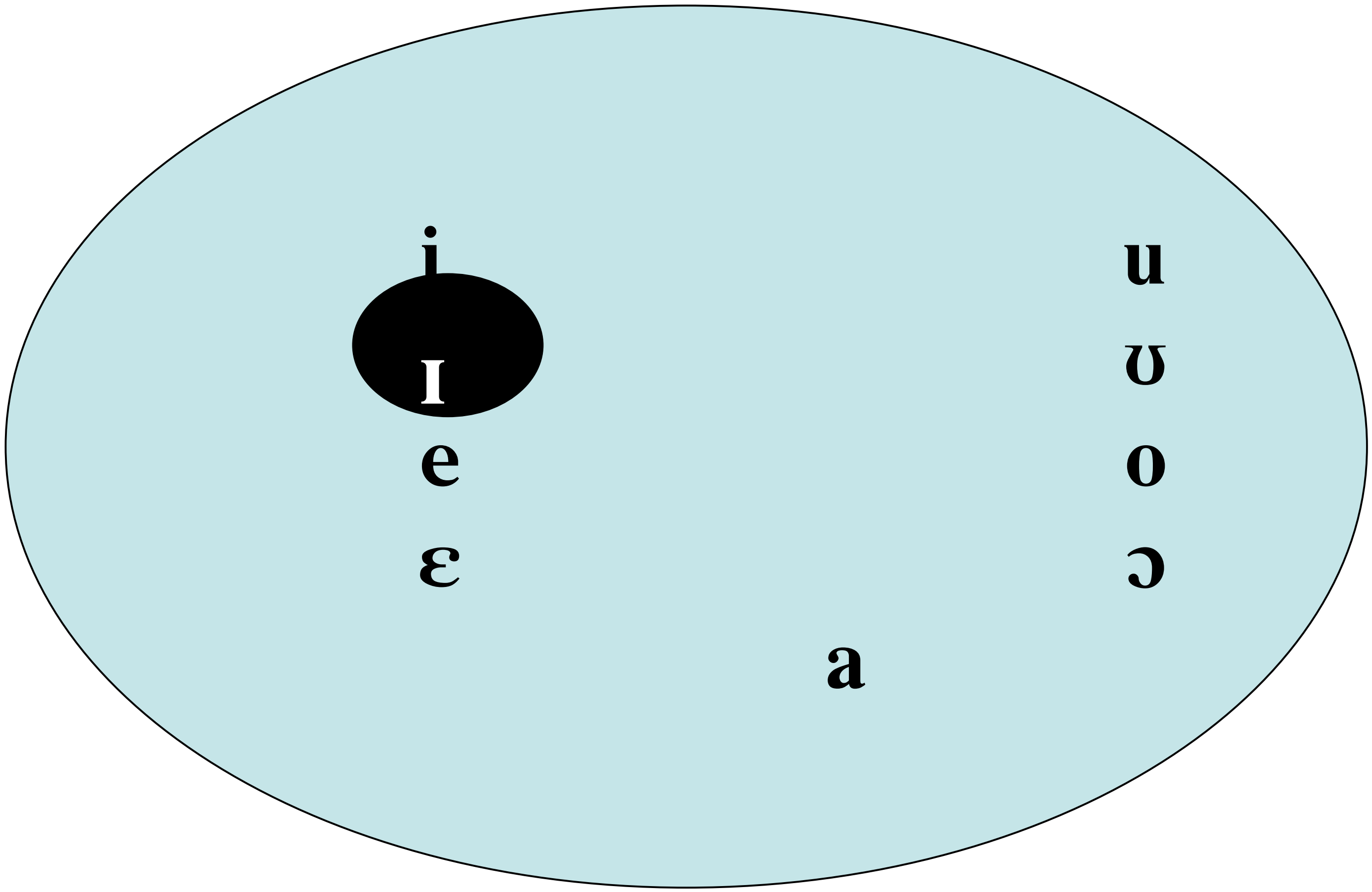
Figure by MIT OpenCourseWare.

Stressed V Duration > 60ms

- French: final stress, except if final V = ə

livrə	Stressed V Duration	Edgemost R
☛ lívrə		*
livrə́	*!	
livre	Stressed V Duration	Edgemost R
lívre		*!
☛ livré		

Another low-faith segment?



Targets of GF in Kinyamwezi

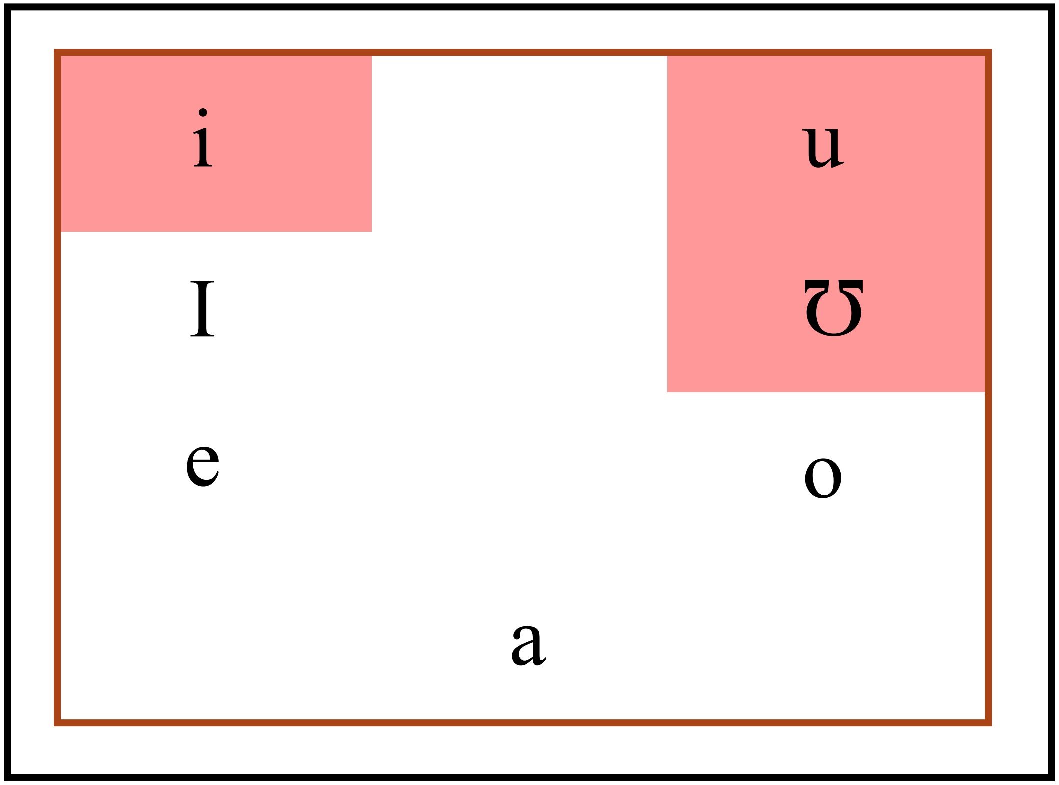


Figure by MIT OpenCourseWare.

Glide formation

i -> j **/mi-áka/ [mjáka]** ‘years’

u -> w **/mu-ana/ [ŋwana]** ‘child’

ʊ -> w **/kʊ-anʊkʊla/ [kwanʊkʊla]** ‘to receive’

ɪ -> V_2 copy

/a-lɪ-anʊkʊla/ [alaanʊkʊla] ‘is receiving’

[-high] V -> V_2 copy

/a-ka-eŋha/ [akeeŋha] ‘she brought’

/…o-V…/ […VV…]

Analysis (ideas from Casali 1997)

- **Markedness**

  ***VV**: V_1V_2, $V_1 \neq V_2$

  Glide: glides are [+high, +ATR]

- **Faithfulness**

  MAX F-root: features of root V's must surface

  MAX [+round]: [+rd] must surface

  DEP [+high], [+ATR]: surface [+hi/ATR] is underlying

- **Ranking of faithfulness**:

  MAX F-root, MAX [+rd], DEP [+high]

  >> DEP [+ATR] >> MAX other

Cʊ-a -> Cwa

Cʊ-a	MAX +rd	MAX F-root	DEP +ATR
Caa	*!		
Coo		*!	
☛ Cwa			*

Co-a -> Caa

Co-a	DEP [+high]	MAX [+rd]
Cwa	*!	
☛ Caa		*

Cɪ-a -> Caa

Cɪ-a	MAX F-root	DEP +ATR	MAX other
Cεε	*!		
Cja		*!	
☛ Caa			*

Cɪ-a -> Cja

Cɪ-a	MAX F-root	DEP +ATR	MAX other
Cεε	*!		
☛ Cja			
Caa			*!

The unnatural class

- **{i, u, ʊ}** comes from the difference between higher and lower ranking faithfulness:
- MAX [+ATR], [+round], MAX root, DEP [+high]

DEP [+ATR]

MAX other

Laterals as ambivalent

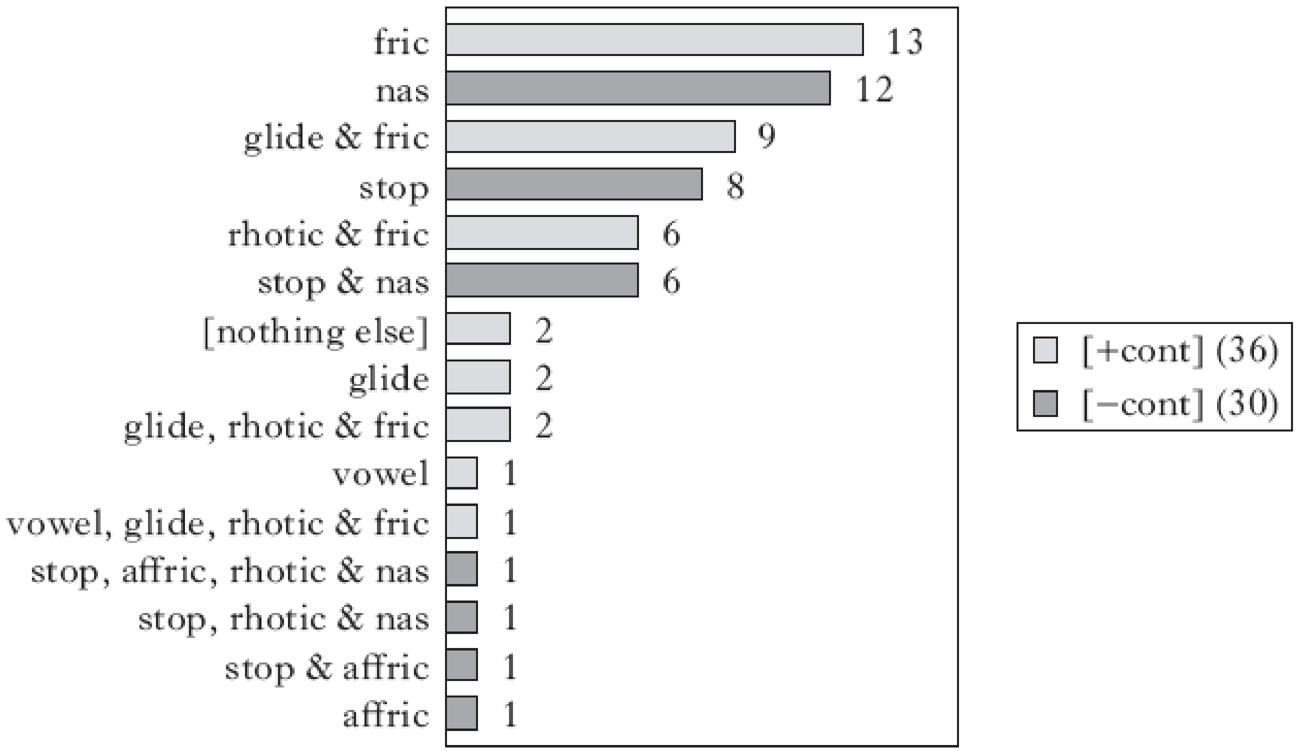


Figure 1

The other members of necessarily [+continuant] and [−continuant] classes containing lateral liquids.

Examples

(8) /l/ *patterning with continuants in Finnish*
(Sulkala & Karjalainen 1992: 387–388)

		active potential	*2nd part active*	
a.	/pur/	[purrut]	[purree]	'bite'
	/nous/	[noussut]	[noussee]	'rise'
	/tul	[tullut]	[tullee]	'come'
b.	/avat/	[avannut]	[avannee]	'open'

(7) /l/ *patterning with non-continuants in Basque* (Hualde 1991: 96)

egu[m] berri	'new day'		
egu[ɱ] fresku	'cool day'		
egu[n̪] denak	'every day'	ata[l̪] denak	'every section'
egu[ɲ] ttiki	'small day'	ata[ʎ] ttiki	'small section'
egu[ŋ] gorri	'red day'		

If this was the right analysis, what would it mean?

- Mielke: continuancy is a continuum, there are no clear boundaries between [+cont] and [-cont]; laterals are located somewhere in the middle, where categorization is hard.
- [This might be true of some features, but why do facts of this sort relate/lend support to the Emergent Feature Theory?]

Questions to ask

- What are the cases are there where one and the same process diagnoses laterals as [+cont] in one language and [-cont] in another? E.g. post-X spirantization where laterals sometimes act like the glides and the fricatives and sometimes like the stops?
- Are the divergent laterals (e.g. Finnish and Basque) identical phonetically wrt degree of oral constriction?
- Try alternative analyses:
  - Maybe Basque says nothing about [l] being [-cont]: we need to block [r] from assimilating, not to explain why [l] assimilates.
  - Maybe Finnish says nothing about [l] being [+cont]: all we need to explain is why [t] doesn't trigger progressive assimilation, not why [l] does.

An unsolved problem
(Flemming 2004)

- Distributional class vs. grammatical class
- Distributional classes are frequently impossible to characterize with a plausible feature set because the distribution results from the interaction of multiple related constraints.
- The problem will be to characterize the property that relates the constraints.

A recurrent unnatural class: Bukusu N deletion

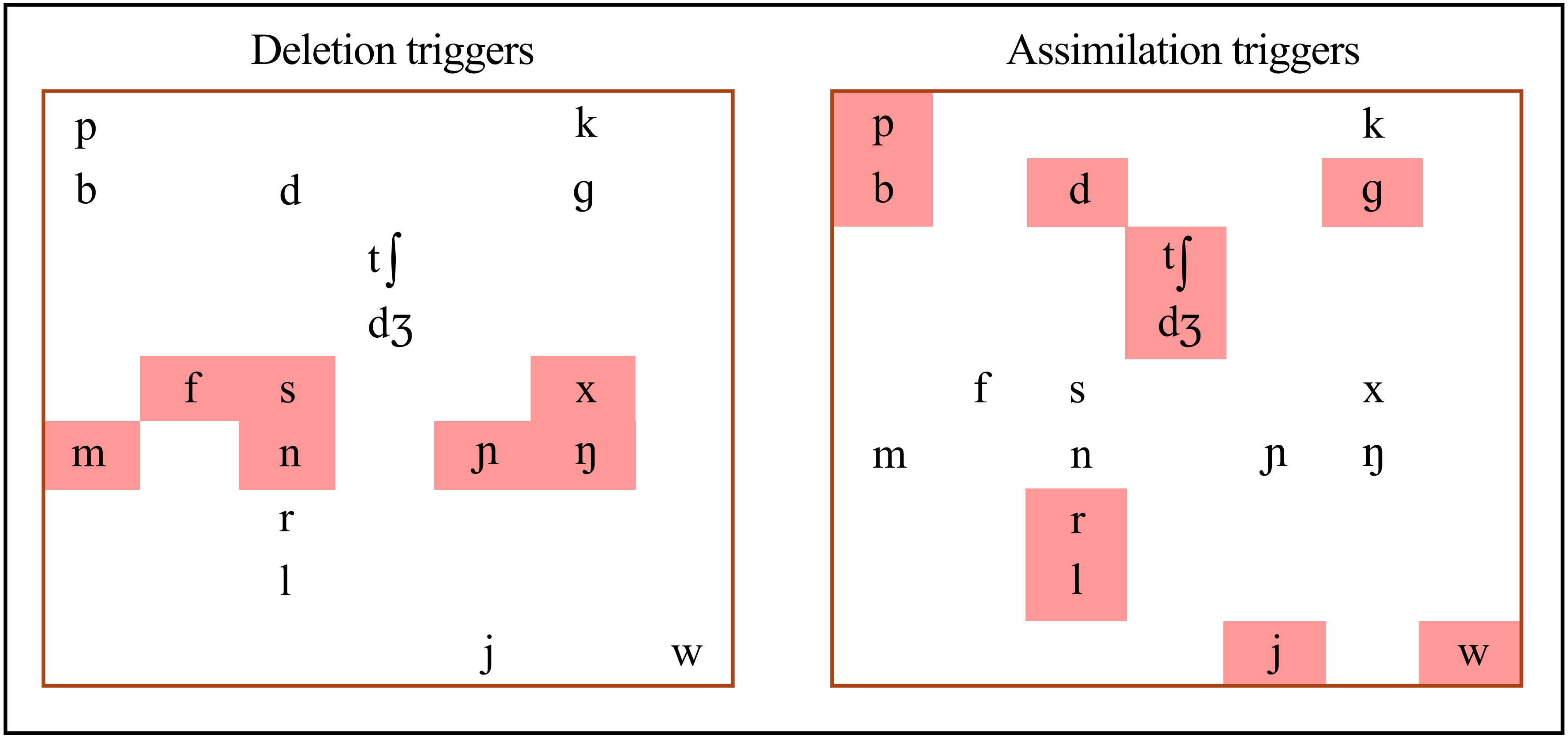


Figure by MIT OpenCourseWare.

N deletion

/i-n-fula/	→	[e:fula]	'rain'
/in-som-ij-a/	→	[e:somia]	'I teach'
/i-n-xele/	→	[e:xele]	'frog'
/in-nuun-a/	→	[e:nuuna]	'I suck'
/in-meel-a/	→	[e:meela]	'I am drunk'
/i-n-ɲaɲa/	→	[e:ɲaɲe]	'tomato'
/i-n-ŋuaŋua/	→	[e:ŋwaŋwa]	'camel'

Figure by MIT OpenCourseWare.

Place assimilation

/in-wulil-a/	→	[embulila]	'I hear'
/in-pim-a/	→	[empima]	'I measure'
/in-bon-a/	→	[embona]	'I see'
/in-ùʃex-a/	→	[eɲèʒexa]	'I laugh'
/ùʃi-n-ɟu/	→	[ùʃiɲèʒu]	'houses'
/ùʃi-n-jim̩b-o/	→	[ùʃiɲim̩bo]	'songs'
/i-n-goxo/	→	[eŋgoxo]	'hen'

Figure by MIT OpenCourseWare.

Cluster of ***NX** constraints
(Flemming 2004)

- Nasal deletion before all of **G, L, N, F**
  Lithuanian
- Or before subsets of these

Hungarian:	**G, L,**	**F (+h)**
Frisian**:**	**G, L,**	**F**
Diola**:**	**G, L**	
Bahasa**:**	**G, L, N**	
Latin**:**	**L**	**F**
Zoque**:**	**L, N**	**F (-h)**

Lithuanian (Kenstowicz via Flemming)

(17) Lithuanian (Kenstowicz, 1972: 12)

(a) No deletion of /n/ before stops.

sáːndora	'covenant'	cf.	dorà	'virtue'
sáːntaka	'confluence'		tekéːti	'to flow'
sáːmbuːris	'assembly'		buːrĩs	'crowd'
sáːmpilas	'stock, store'		pìlnas	'full'
sáːmbuːris	'assembly'		buːrĩːs	'crowd'
sáːŋkaba	'coupling, clamp'		kãːbeː	'hook'

(b) Deletion of /n/ before glides, fricatives, liquids and nasals.

sáːjuŋga	'union'	cf.	jùŋgas	'yoke'
sáːvoka	'idea'		vóːkti	'understand'
sáːskambis	'harmony'		skambéːti	'ring'
sáːʃlavos	'sweepings'		ʃlúoti	'sweep'
sáːʒineː	'conscience'		ʒinóːti	'know'
sáːliːtis	'clash, contact'		líːti	'to rain'
sáːraʃas	'list, register		raʃíːti	'to write'
sáːmokslas	'conspiracy'		móːkslas	'skill'
sáːnariːs	'joint'		narĩːs	'link'

Different *NX constraints

- Can't normally be diagnosed in one language via interaction with faithfulness

  *NL >> Faith 1 >> *NS >> Faith 2, normally same as

  *N/{_L, S} >> Faith 1 >> Faith 2
- Can be identified in different morphological context

  [*NG, *NL, *NS stem-internally in Latin

  *NL, *NS, not *NG active at prefix-stem boundaries]

Bukusu

- could be one of these:

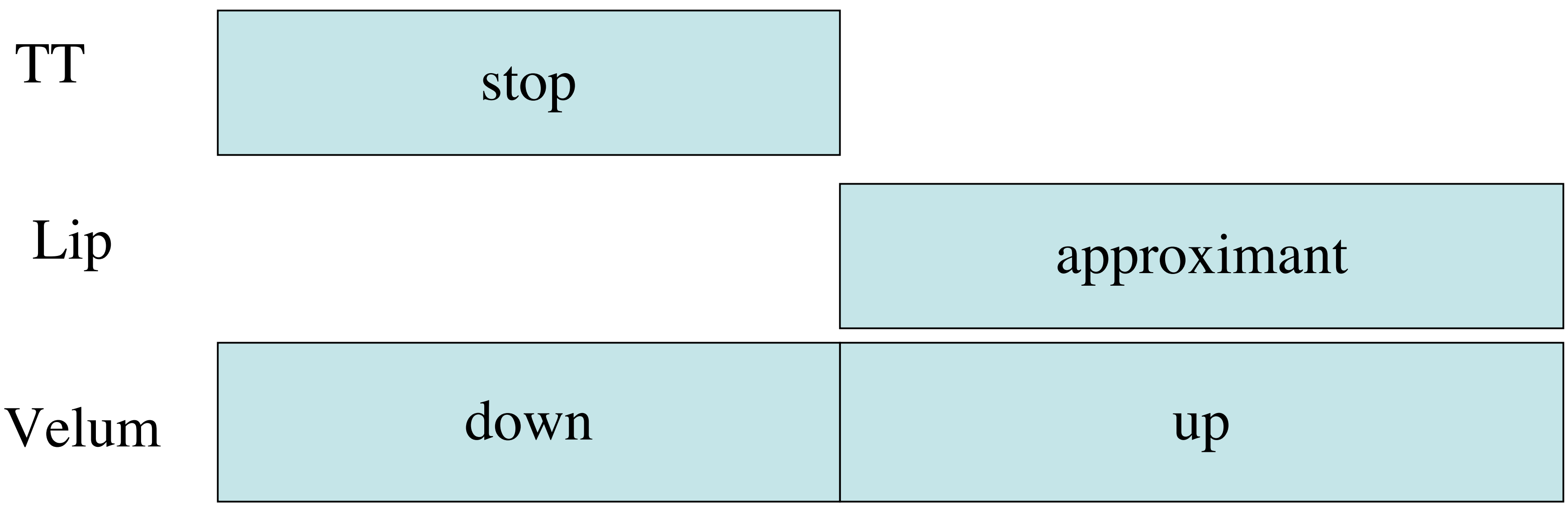
• Bukusu	**N**	**F**
Hungarian:	**G, L,**	**F (+h)**
Frisian**:**	**G, L,**	**F**
Diola**:**	**G, L**	
Bahasa**:**	**G, L, N**	
Latin**:**	**L**	**F**
Zoque**:**	**L, N**	**F (-h)**

“the class of segments that conditions deletion of a preceding nasal is derived from the combined action of all of the constraints that are ranked above [faithfulness]; a set of sounds can also form a natural class if they are **marked in the same environment** according to multiple constraints […] it is this property of being marked in the same environment that makes sets of sounds a possible natural class.[…] the range of possible natural classes is derived from the nature of the constraint set, not from the nature of the feature set.”

(Flemming, p. 23 in ms. version)

Why N's are marked before G,L, F

- Place assimilation (more likely for nasals than oral C's)
- CC-Place assimilation = extension of oral constriction
- Transmits: constriction degree/mode, [contin], [cons], [lateral]
- Browman&Goldstein 1992 Phonetica; Padgett 1994 NLLT
- nw -> w̃w -> w

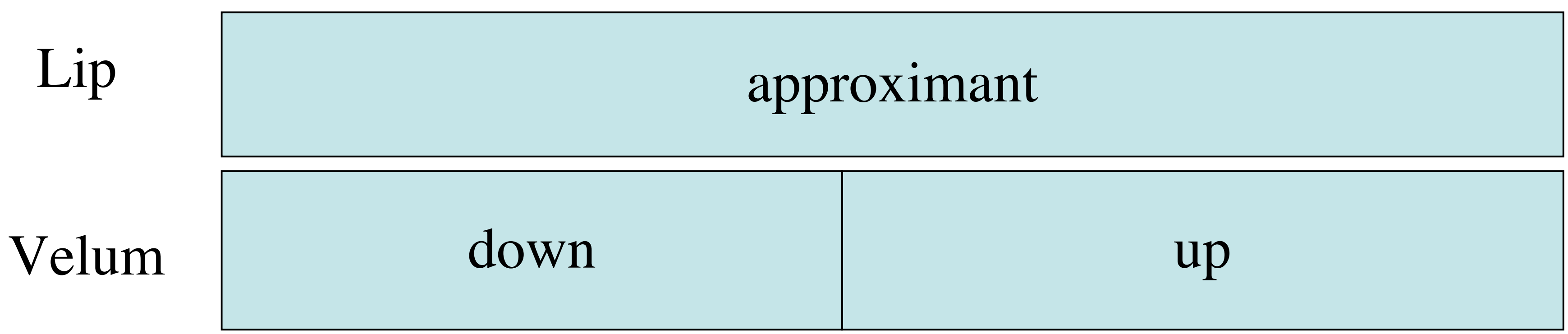


"the class of segments that conditions deletion of a preceding nasal is derived from the combined action of all of the constraints that are ranked above [faithfulness]; a set of sounds can also form a natural class if they are **marked in the same environment** according to multiple constraints […] it is this property of being marked in the same environment that makes sets of sounds a possible natural class.[…] the range of possible natural classes is derived from the nature of the constraint set, not from the nature of the feature set."

(Flemming, p. 23 in ms. version)

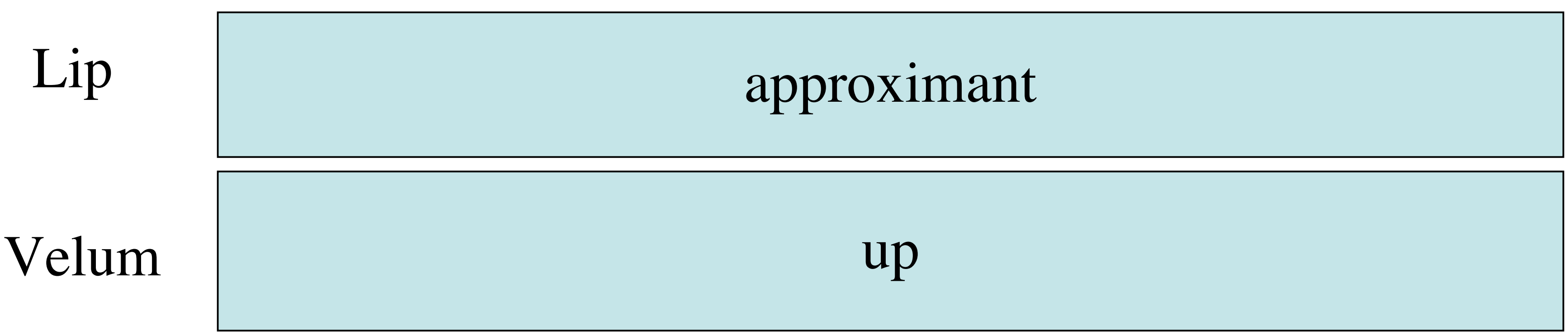
Why N’s are marked before G,L,N,F

- Place assimilation (more likely for nasals than oral C’s)
- CC-Place assimilation = extension of oral constriction
- Transmits: continuant, consonantal, lateral
- Browman&Goldstein 1992 Phonetica; Padgett 1994 NLLT
- nw -> w̃w -> w



Why N’s are marked before G,L,N,F

- Place assimilation (more likely for nasals than oral C’s)
- CC-Place assimilation = extension of oral constriction
- Transmits: continuant, consonantal, lateral
- Browman&Goldstein 1992 Phonetica; Padgett 1994 NLLT
- nw -> w̃w -> w



Nasal deletion before {G, L, N, F} linked to place assimilation

- N's delete where they place-assimilate: *before* _C
- N's less likely to place-assimilate are less likely to delete:
- Lithuanian **m**

  Unassimilated before stops: **gim-ti** 'be born-inf'

  Undeleted before F, N, L, G: **krimʃ-ti** 'chew-inf'
- Lithuanian **n**

  Assimilated before stops: **sa[m]-buris** 'assembly'

  Deleted before F, N, L, G: **sã-ʒine** 'consciousness'

Latin

- **m** is unassimilated, undeleted before F:

  em-(p)tum ‘bought-ppl’

  em-(p)si ‘bought-1sg’

- **n** is assimilated, deleted before F:

  in-potens -> [impotens]

  in-superabilis -> [i):superabilis]

Zoque (Wonderly 1951)

- Nasal deleted where subject to assimilation: #_C

  kwarto, ŋgwarto ‘room; my room’

  faha, faha ‘belt, my belt’

  sʌk, sʌk ‘beans, my beans’

  ʃapun, ʃapun ‘soap, my soap’

- Nasal medially not assimilated _T, or deleted _F:

  tʃoʔngoja ‘rabbit’ **tototuʔnba** ‘he reads’

  ʔaŋsʌŋ ‘season’, **ʔaŋham** ‘to feel’

  pʌnʃehmah ‘still like a man’

There is no *NN

	G	L	N	F	N:?
Bahasa	√	√	√	√	*
Zoque		√	√	√	*(#_)
Bukusu			√	√	*
Lithuanian	√	√	√	√	*
Hungarian	√	√		√	ok
Latin		√		√	ok
Diola	√	√			ok
Frisian	√	√		√	?

Interaction of ***NX** and ***N**:

- Where geminate **N:** is allowed no ***NN** effect.
- Suggests that ***NX** is triggered not directly by adjacency of N to X but by place assimilation of N to X, generating a nasalized X. When X is a nasal and place assimilation yields a geminate N:, output should surface as such, if N: is allowable generally.
- Confirms that ***NX** is not a unified constraint.
- Confirms that the proper set of constraints is
- ***nasalized F, *nasalized L, *nasalized G**
- Independent evidence for these: nasal harmony

The question

- If the scenario **nw -> w̃w -> w** is right then the constraint is not ***NX** but ***nasalized X**, and **X** refers to a subset of non-stops.
- But what do the **X** segments share? It's not an accident that all avoid being nasalized even if some avoid it more than others.

  [cf. "a set of sounds can also form a natural class if they are **marked in the same environment** according to multiple constraints".]
- Maybe what this shows is that we don't have all the features we need.

Another open question: *g

- Some processes always target the same natural classes: e.g. postnasal voicing always targets all stops, regardless of p.o.a.
- If constraints against individual voiced stops exist, e.g. *g, systems are predicted in which the ranking *g >> *NT blocks postnasal voicing from applying to some stops.
- Flemming: disallow context free *g; [g] might be bad in some contexts but will never be bad post-N.
- And Arabic? All context sensitive anti-[g] constraints have conspired to rise to the top?