

Class 13: Markedness and allomorph choice

Outline

- Phonologically based allomorph selection
- Allomorph selection as TETU
- Case studies
- Morphological ineffability

1. Phonologically based allomorph selection

French

(based on discussion in Joan Mascaró (1996) External allomorphy and contractions in Romance, *Probus* 8, 181-205)

‘friend’	‘husband’	
[bɛl] ami	[bo] mari	pretty, nice
[nuvɛl] ami	[nuvo] mari	new
[vjɛj] ami	[vjø] mari	old
[sɛt] ami	[sø] mari	stupid
[kɛl] ami	[kɛl] mari	what
[ʒɔli] ami	[ʒɔli] mari	pretty

Despite the regularities in the correspondences between the two allomorphs, we wouldn’t want to try to derive them both from a single UR.

Rather, both are probably listed, but the choice depends on phonological factors (whether the following word begins with a consonant or a vowel).

Korean

	‘baby’	‘water’
subject	ko.gi-.ga	mu.R-i
object	ko.gi-.ɾɿl	mu.R-ɿl
topic	ko.gi-.nɿn	mu.R-ɿn
‘it’s ...’	ko.gi-.je.jo	mu.R-i.e.jo

Other examples

- English *a/an*, [ðʌ]/[ði]
- Italian *il/l’*
- English *more X/Xer*
- English deverbial-noun *-al/-ment*

2. Allomorph selection as TETU

The key proposal of OT is that there can be output-oriented constraints that are violable, but nonetheless active.

Given a theory (or at least a partial inventory) of markedness conditions, we expect them to pop up here and there, even in languages that seem to violate them rather freely.

Reduplicative TETU: CORR-IO >> PHONO >> CORR-BR

Allomorphic TETU: CORR-IO >> PHONO, but the lexicon sometimes supplies an allomorph that satisfies PHONO:

(again, from Mascaró)

{/bɛl/, /bo/} ami	MAX-C	DEP-C	NoCODA	ONSET
→ bɛl ami				
bo ami				*!

/ʒɔli/ ami	MAX-C	DEP-C	NoCODA	ONSET
ʒɔlit ami		*!		
→ ʒɔli ami				*

{/bɛl/, /bo/} maɾi	MAX-C	DEP-C	NoCODA	ONSET
bɛl maɾi			*!	
→ bo maɾi				

/kɛl/ maɾi	MAX-C	DEP-C	NoCODA	ONSET
→ kɛl maɾi			*	
ko maɾi	*!			

Moral: even though French has no repair mechanism for hiatus or codas, it nevertheless avoids them when it can do so at no cost to faithfulness.

3. Case study: Tagalog nasal substitution

Nasal substitution

a.	<i>h</i>	hukbó	‘army’	paŋ-hukbó	‘military’
	<i>m</i>	marká	‘mark’	paŋ-marká	‘marker’
		(no examples of <i>n</i>)			
	<i>ŋ</i>	ŋálit	‘grinding of teeth’	paŋ-ŋa-ŋálit	‘grinding of teeth’
	<i>w</i>	mag-wisík	‘to sprinkle’	paŋ-wisík	‘sprinkler’
	<i>j</i>	jamót	‘annoyance’	maŋ-jamót	‘to annoy’
b.	<i>l</i>	lágom	‘assimilation’	ma-pan-lágom	‘monopolistic’
c.	<i>r</i>	rasjón	‘ration’	paŋ-rasjón, pan-rasjón	‘for rationing’
d.	<i>p</i>	pighatí?	‘grief’	pa-mi-mighatí?	‘being in grief’
		poʔók	‘district’	pam-poʔók	‘local’
	<i>t</i>	pag-tú:loj	‘staying as guest’	ka:-pa-nulú:j-an	‘fellow lodger’
		tabój	‘driving forward’	pan-tabój	‘to goad’

<i>s</i>	sú:lat	‘writing’	ma:-nu-nulát	‘writer’
	sú:lat	‘writing’	pan-sú:lat	‘writing instrument’
<i>k</i>	kamkám	‘usurpation’	ma-pa-ŋamkám	‘rapacious’
	kaliskís	‘scales’	paŋ-kaliskís	‘tool for removing scales’
<i>ʔ</i>	ʔisdaʔ	‘fish’	ma:-ŋi-ŋisdáʔ	‘fisher’
	ʔulól	‘silly’	maŋ-ʔulól	‘to fool someone’
<i>b</i>	mag-bigáj	‘to give’	ma-migáj	‘to distribute’
	bigkás	‘pronouncing’	mam-bi-bigkás	‘reciter’
<i>d</i>	dalá:ŋin	‘prayer’	ʔi-pa-nalá:ŋ-in	‘to pray’
	diníg	‘audible’	pan-diníg	‘sense of hearing’
<i>g</i>	gindáj	‘unsteadiness on feet’	pa-ŋi-ŋindáj	‘unsteadiness on feet’
	gá:waj	‘witchcraft’	maŋ-ga-gá:waj	‘witch’

4. PHONOS

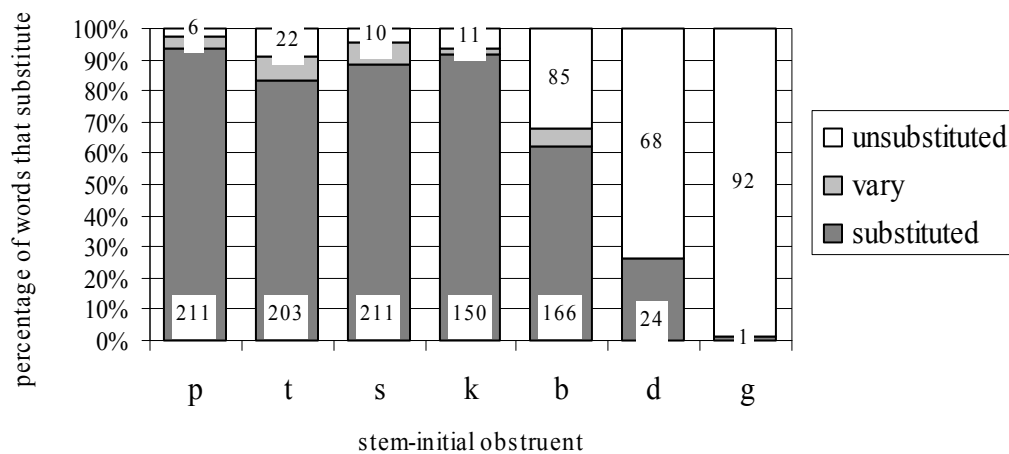
Tagalog freely tolerates *mp*, *nt*, *ns*, *ŋk*

*NC_o >> {IDENT-IO[VOICE], MAX-C, DEP-V, UNIFORMITY, IDENT-IO[NASAL], etc.}

Tagalog also tolerates onset *ŋ* (cf. Japanese, English).

*[ŋ] >> {IDENT-IO[PLACE], MAX-C, IDENT-IO[NASAL], etc.}

5. Distribution of nasal substitution

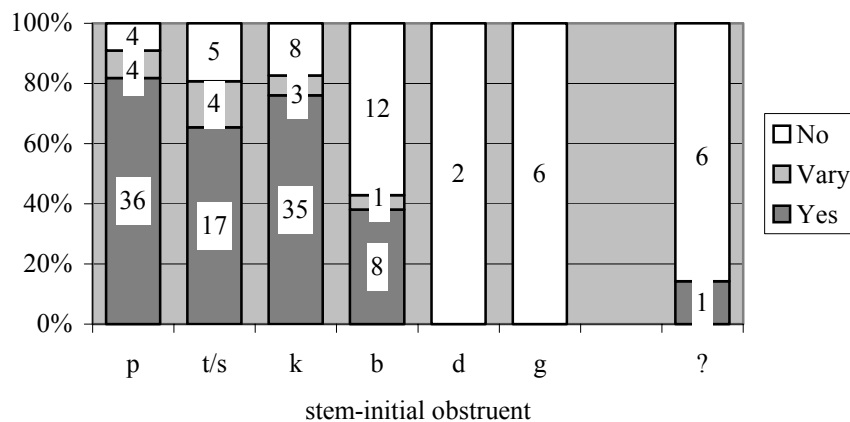


- Nasal substitution is frequent when it eliminates a violation of *NC_o
- Nasal substitution is infrequent when it creates a violation of *[ŋ]

(actually, I propose a scale *[ŋ] >> *[n] >> *[m]; the phonetic property that presumably makes [ŋ] a bad onset is that the oral “side tube” is very short, so there are no low-enough antiformants to interfere with the vowel-like-ness of the sound; [n] has a slightly longer oral tube, but the lowest antiformant is still higher than [m]’s)

6. Loans

This distribution is fairly productive—if you can say that of a distribution:



7. Cross-linguistic facts

(adapted from Newman 1984, p. 10)

	substituted?						ranking
	p	t,s	k	b	d	g	
Toba Batak type	+	~	-	~	-	-	*[ŋ] >> *NC ₀ >> {MORPHUNI, *[n, *[m] >> NASSUB
Malay type	+	+	+	-	-	-	*NC ₀ >> {MORPHUNI, *[ŋ, *[n, *[m] >> NASSUB
Sama Badjao type	+	+	+	+	-	-	*NC ₀ >> {*[ŋ, *[n] >> NASSUB >> {*[m, MORPHUNI}
Cebuano type	+	+	+	+	+	-	*NC ₀ >> *[ŋ] >> NASSUB >> {*[n, *[m, MORPHUNI}
Kalinga type	+	+	+	+	+	+	NASSUB >> {*[ŋ, *[n, *[m, MORPHUNI}

(Caution: Tagalog is often described as being a Sama-Badjao- or Cebuano-type language, though the facts are more complicated; the same may be true for some other languages. You will see another interesting pattern in Timugon Murut in the next problem set.)

8. Case study: Finnish genitives

Anttila, Arto. 1997. Deriving variation from grammar. In Frans Hinskens, Roeland van Hout, and Leo Wetzels (eds.), *Variation, Change and Phonological Theory*, 35-68. Amsterdam/Philadelphia, John Benjamins Publishing Company.

9. Distribution of Finnish genitives

Some stems always take the ‘strong’ genitive *-iden*, some always take the ‘weak’ genitive *-en/-jen*, and some vary, but often with a preference one way or the other.

- Monosyllables always take the strong variant (INITIALSTRESS, *STRESSEDLIGHT)
/maa/ má-i.den *má.-jen

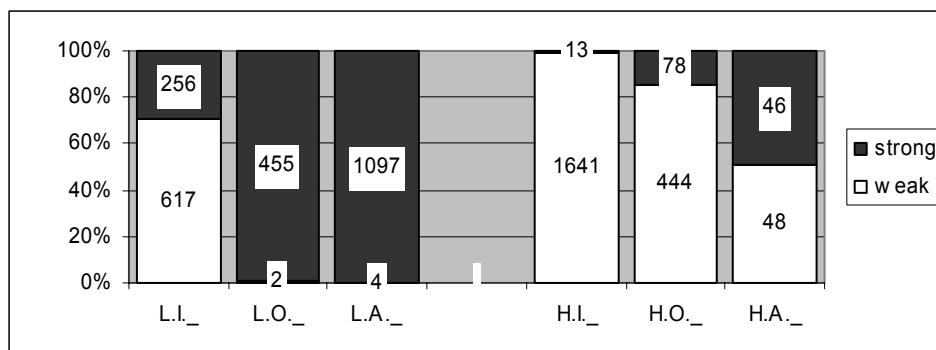
- Disyllabic stems ending in a light syllable always take the weak variant (INITIALSTRESS, *STRESSCLASH, *UNSTRESSEDHEAVY)
/kala/ *ká.lo-i.den, *ká.lò-i.den ká.lo.-jen
- Disyllabic and longer stems ending in a heavy syllable always take the strong variant (IDENTWEIGHT?? These cases aren't really discussed.)
/pálttoo/ pált.to-i.den *pált.to.-jen
- Trisyllabic and longer stems ending in a light syllable vary.
 - Those ending in a high vowel prefer the weak variant
/lemmikki/ ~lém.mik.kè-i.den lém.mik.ki.-en
 - Those ending in a low vowel prefer the strong variant
/sairaala/ sái.raa.lò-i.den ~sái.raa.lo.-jen
 - Those ending in a mid vowel vary more freely (secondary stress is optional: *LAPSE must be freely ranked w.r.t. some anti-stress constraint)
/fyysikko/ fyy.si.kò-i.den fyy.sik.ko.-jen

Weird quirk: these generalizations refer to *underlying* vowel height

*HEAVYHIGH >> *HEAVYMID >> *HEAVYLOW
 *LIGHTLOW >> *LIGHTMID >> *LIGHTHIGH
 *STRESSEDHIGH >> *STRESSED MID >> *STRESSEDLOW
 *UNSTRESSEDLOW >> *UNSTRESSED MID >> *UNSTRESSEDHIGH
 (do we really need all four scales?)

- In trisyllabic and longer stems, there's also a tendency for a heavy antepenult to take the weak genitive and for a light antepenult to take the strong genitive. ("weight-clash/lapse" constraints: *H.H, *L.L)

Weight-clash considerations conflict with vowel-height considerations (corpus data reported by Anttila, for 3-, 4- and 5-syllable words combined):

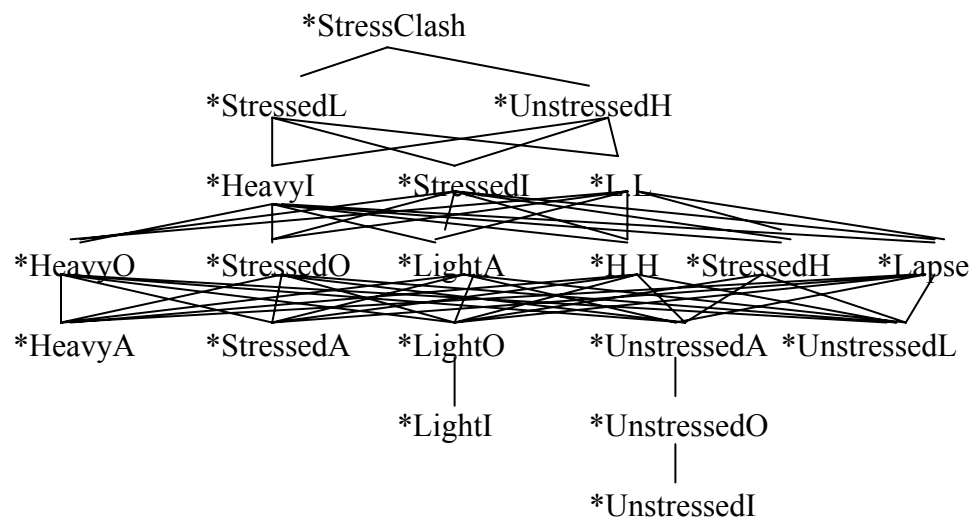


Some categorical gaps in longer words:

/ministeri/	mí.nis.te.rè-i.den	mí.nis.te.ri.-en
/margariini/	*már.ga.rii.ne-i.den	már.ga.rii.ni.-en
	*már.ga.rii.nè-i.den	
	*már.ga.rii.nè-i.den	
/aleksanteri/	á.lek.sàn.te.rè-i.den	á.lek.sàn.te.ri.-en
/koordinaatisto/	kóor.di.nàa.tis.tò-i.den	kòor.di.nàa.tis.to.-jen
/italiaano/	*í.ta.li.àa.no-i.den	í.ta.li.àa.no.-jen
	*í.ta.li.aa.nò-i.den	
	*í.ta.li.àa.nò-i.den	

10. Proposed grammar

Nearly stratal, but not quite



(plus transitivity)

There is variation in ranking within each row, and among all the constraints of the last three rows, except that *LIGHTO >> *LIGHTI and *UNSTRESSED A >> *UNSTRESSED O >> *UNSTRESSED I.

How does this work in terms of statistics? (discuss)

11. Case study: French *-esque*

Marc Plénat (1997). Analyse morpho-phonologique d'un corpus d'adjectifs dérivés en *-esque*. *French Language Studies* 7, 163-179.

-esque

Pretty much like English *-esque*.

Forms adjective from noun (common or proper) or noun phrase; very productive.

buanderie	'laundromat'	buanderesque
Clinton		clintonesque
tour de force	'feat'	tourdeforcesque

Plénat investigates a mostly print corpus of about 800 adjectives in *-esque*, including many nonce formations.

12. Latent final consonants

Famous feature of French: many underlying word-final consonants are reflected in the orthography but pronounced only when pre-vocalic, including before a silent /œ/.

trop	[tʁo]	'too'
trop acide	[tʁopasid]	'too sour'
charmant	[ʃaʁmã]	'charming (masc.)'
charmante	[ʃaʁmãt]	'charming (fem.)'

These consonants generally show up in the *-esque* forms:

Fragonar(d)	fragonar[d]-esque
Danto(n)	danto[n]-esque

Surprisingly, orthography alone is not a sufficient guide; there are some unexpected forms:

Hersan(t)	Hersan[t]-esque	or	Hersan_-esque
tobogga(n)	tobogga[n]-esque	or	tobogga(n)[t]-esque

13. Overt final consonants

Generally preserved

plantigra[d]	plantigra[d]-esque
Clinto[n]	clinto[n]-esque

14. Consonantal OCP effects: *sVs (or *TAUTOSYLLABIC SIBILANTS)

Except in monosyllabic stems or last-word-of-stems (far[s], far[s]-esque),

V s → Ø / X __ -esque where X ≠ s
[-nas]

Cervante[s] cervant-esque
juliéna[s] julién-esque
CNRS ([seeneʁes]) CNR-esque

s → Ø / s V __ -esque
[-nas]

Onassi[s] Onassi-esque

V s → Vn / X __ -esque where X ≠ s
[+nas]

Camo[ãs] camo[in]-esque

Let's try to analyze:

far[s]+esque	STEM ≥ 1 SYLL	*{A,E} V	*sVs	MAX-C	*IV	MAX-V
☞ far[s]-esque			*			
fa-esque		*!		**		
f-esque	*!			**		*

(Why not *far-esque? Maybe coda /r/ doesn't want to become onset?)

Cervant[es]+esque	STEM ≥ 1 SYLL	*{A,E} V	*sVs	MAX-C	*IV	MAX-V
cervant[es]-esque			*!			
cervant[e]-esque		*!		*		
☞ cervant-esque						*

muliebr[is]+esque	STEM ≥ 1 SYLL	*{A,E} V	*sVs	MAX-C	*IV	MAX-V
muliebr[is]-esque			*!			
muliebr[i]-esque				*	*!	
☞ muliebr-esque				*		*

Ona[s]i[s]+esque	STEM \geq 1 SYLL	*{A,E} V	*sVs	MAX-C	*IV	MAX-V
Ona[sis]-esque			*!*			
☞ Ona[si]-esque				*	*	
Onas-esque			*!	*		*
Ona-esque		*!		**		*

Camo[ãs]+esque	STEM \geq 1 SYLL	*{A,E} V	*sVs	MAX-C	*IV	MAX-V	INTEG
camo[ãs]-esque			*!				
camo[ã]-esque		*!		*			
☞ camo[an]-esque				*			*
camo-esque		*!		*		*	
cam-esque				*		*!*	

Latent /s/ and /z/ behave similarly: they don't show up in stems longer than 1 syllable, and the preceding vowel usually disappears (except one example of [u]):

clapoti(s) clapot-esque

In monosyllables they may show up, or a different C may appear:

Louis II(z) louisdeu[z]-esque
gueule de boi(z) gueuledeboi[t]-esque

Variable ranking between *{s,z} Vs and DEP?¹

No examples like *Onassis*, but with final latent /s/, though.

15. Consonantal OCP effects: *KV(s)K (or *TAUTOSYLLABICVELARS)

In some velar-final words, nothing happens:

ga[g] ga[g]-esque
Pétrar[k] pétrar[k]-esque

In others, the suffix changes:

Mar[k] mar[k]-este
Silvio Pelli[k]o sylviopelli[k]-este

In some stems longer than one syllable, the velar disappears

panégryri[k] panégyr-esque
Goeri[ŋg] goeri[n]-esque

¹ A note on epenthetic consonants: Plénat notes that many have a precedent in another allomorph: *fer blanc*, *ferblantesque* (cf. *ferblantier*). Reminiscent of Steriadean lexical conservatism.

Mar[k]+esque	STEM ≥1 SYLL	*{A,E}V	IDENTPLACE STEM	*KV(s)K	IDENTPLACE AFFIX	MAX-C	*IV	MAX-V
☞ Mar[k]-esque				*				
Ma-esque		*!				**		
M-esque	*!					**		*
☞ Mar[k]-este					*			
Mart-esque			*!					

panégyr[ik]+esque	STEM ≥1 SYLL	*{A,E}V	IDENTPLACE STEM	*KV(s)K	IDENTPLACE AFFIX	MAX-C	*IV	MAX-V
☞ panégyr[ik]-esque				*				
panégyr[i]-esque						*	*!	
☞ panégyr-esque						*		*
☞ panégyr[ik]-este					*			
panégyr[it]-esque			*!					

16. Consonantal OCP effects: *[-son]V[-son]

Loss of final vowel is less likely if the preceding consonant is an obstruent (significantly so in words with 3 syllables or more: 78% vs. 94%)

17. Vocalic OCP effects: *EC₀E

General French fact: When a latent nasal is restored after [ẽ], or a latent oral C after [e], the vowel becomes [ɛ] (or maybe sometimes [œ] or [e], depending on the suffix):

<i>masc.</i>	<i>fem.</i>
olympi[ẽ](n)	olympi[ɛn]
coutumi[e](r)	coutumi[ɛʁ]
cad[e](t)	cade[ɛt]

When this would happen with *-esque* ([ɛsk]), the vowel is deleted, rather than the consonant restored:

olympi[ẽ](n)	olympi-esque
coutumi[e](r)	coutumi-esque
cassoul[e](t)	cassoul-esque

(only four examples of 2-syllable stems; 2 fail to delete the vowel)

This doesn't happen with overt consonants, except for one example (*Nibelungen*):

ant[ɛn]	ant[ɛn]-esque
coccin[ɛl]	coccin[ɛl]-esque

coccin[ɛl]+esque	STEM ≥1 SYLL	*{A,E}V	MAX-C	*IV	*EC ₀ E	MAX-V
☞ coccin[ɛl]-esque					*	
coccin[ɛ]-esque		*!	*		*	
coccin-esque			*!			*

cassoul[e](t)+esque	STEM ≥1 SYLL	*{A,E}V	MAX-C	*IV	*EC ₀ E	MAX-V
cassoul[et]-esque					*!	
cassoul[e]-esque		*!			*	
☞ cassoul-esque						*

But here we have a contradiction:

coutumi[e](r)+esque	STEM ≥1 SYLL	*{A,E}V	MAX-C	*IV	*EC ₀ E	MAX-V
coutumi[er]-esque					*!	
coutumi[e]-esque		*!			*	
☹ coutumi-esque				*!		*
coutum-esque						**

We saw before that *IV >> MAX-V, because *panégyr-esque* (from *panégyr[ik]*) defeats **panégyr[i]-esque*.

Perhaps we should treat /i/ that becomes nuclear [i] in the stem differently from /i/ that becomes [j] in the stem:

coutumi[e](r)+esque	STEM ≥1 SYLL	*{A,E}V	MAX-C	*IV	*EC ₀ E	MAX-V
coutumi[er]-esque					*!	
coutumi[e]-esque		*!			*	
☞ coutumi-esque				*!		*
coutum-esque			*!			**

Loss of a final V is also less likely (though not significantly) when the penultimate V is part of the “E” family (63% vs. 76%).

And, some ultima “E”s change quality, if there’s a precedent in another allomorph:

not[ɛʁ] not[aʁ]-esque cf. not[aʁ]-iat

18. Hiatus avoidance: *VV family

Vowel-final bases may retain or lose the final vowel, or there can be epenthesis

Fellin[i] fellin[i]-esque or fellin-esque

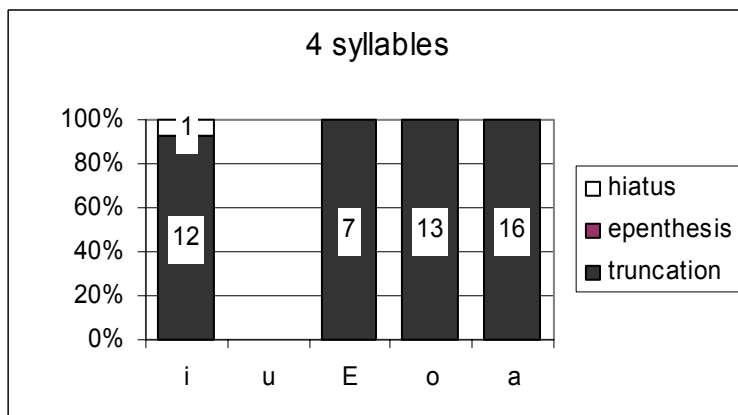
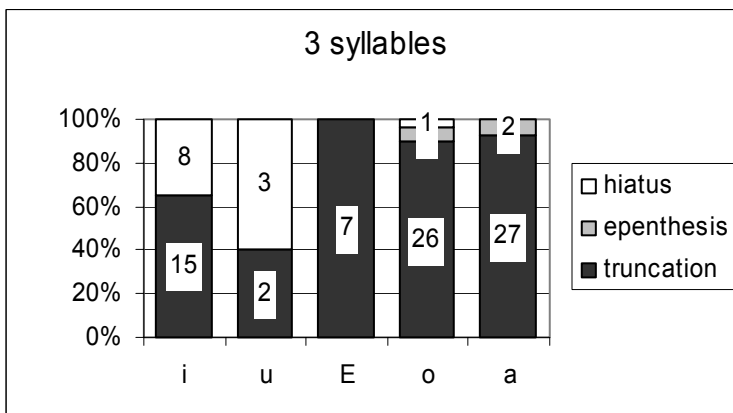
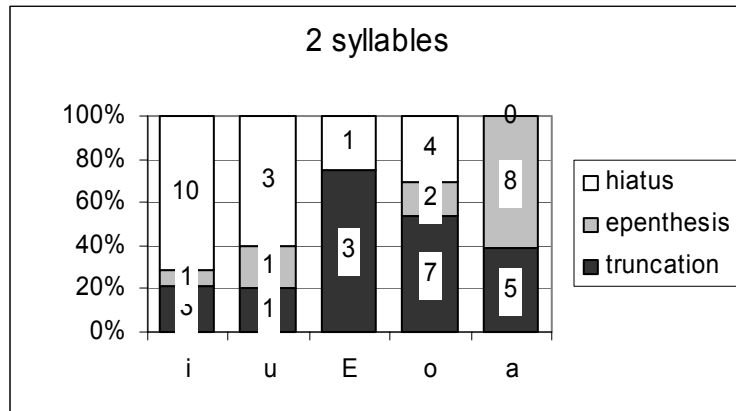
Coct[o] coct[o]-esque or coct-esque

Goy[a] goy[at]-esque

Hiatus is most frequent with higher vowels, and truncation is most frequent with longer words (see graphs).

*AV >> *EV >> *IV (hiatus is less bad the more glide-like the first V is?)

Truncation is less bad the bigger a percentage of the stem you keep????



19. Morphological ineffability

Extension that was last year being contemplated (I don't know what he thinks about this these days) by Bruce Hayes to his recent proposal about componentiality in metrics (ms.).

Ill-formed meter

Why does an ill-formed verse or line not always suggest its own repair?

Hayes's example:

well-formed as iambic pentameter, not complex:

The li- / on dy- / ing thrust- / eth forth / his paw (Shakespeare, R3)

well-formed as iambic pentameter, very complex:

Let me / not to / the mar- / riage of / true minds (Shakespeare, sonnet 116)

ill-formed as iambic pentameter:

Ode to / the West / Wind by / Percy / Bysshe Shelley (Halle & Keyser 1971)

When a line doesn't work, you just have give up and try a different way of saying what you want to say—the grammar doesn't tell you what to do.

20. Ill-formed morphology

Raffelsiefen, Renate (1998), Phonological constraints in English word-formation. In: G. Booij und J. van Marle (Hrsg.) Yearbook of Morphology 1998. Dordrecht: Kluwer.

Many findings, including deadjectival-verb-forming *-en* can't attach to sonorant-final stems:

blacken	*greenen
whiten	*bluen
redde	*brownen
thicken	*thinnen

But **greenen* doesn't suggest a phonological repair. You just have to try again: *greenify?* *greenv?* *make green?*

21. Componentiality in metrics

Metrical component (ranking depends on meter in use) ranks certain "metrical" constraints (the ones inviolable in that meter) higher than "phonological" ones—i.e., it does suggest a repair.

Hayes's example: **Young Emily in her cha----mber* (illegal in "4343" verse)

"4343" grammar says *LAPSE >> DEP-SYLL

⇒ line is repaired as *Young Emily in her chambeler* (or some such). Hayes calls this the “suicide candidate”.

The paraphonological component (“para” because different from phonology of ordinary speech), however, while it allows some outputs that are illegal in colloquial speech (e.g. *o’er*), does not allow **Young Emily in her chambeler*.

Because there is no common legal output of the two components, the derivation crashes, and the line is unmetrical.

22. Componentiality in morphology?

The morphological component absolutely requires that *–en* be attached to an obstruent-initial stem. Thus, we might have suicide candidates like *greenden*, *greeden*:

green+en	MORPHO CONSTRAINTS	PHONO CONSTRAINTS
greenen	*!	
☞ greenden		*
☞ greeden		*

The phonological component, though it allows some optionality in, say, release of final stops, does not allow insertion or denasalization in this context:

green+en	DEP	IDENT[NAS]
☞ greenen		
greenden	*!	
greeden		*!

Because there’s no shared output, the derivation crashes.

23. French, Finnish, Tagalog?

In the cases that we’ve seen, the affixes can attach to just about anything (i.e., the morphological component is vacuous, having no morphophonological constraints relevant to the affix in question), and the phonological constraints do all the work of deciding how the result is pronounced.

It would be nice to find an affix that is phonologically selective in what it will attach to (i.e., there is some ineffability), and also displays phonologically driven allomorphy when it does attach... Can you think of any?

Next: Exemplars and neighborhoods

To do: Assignment #5 (Timugon Murut), start reading Pierrehumbert

