Linguistics 201A Winter 2023
Phonological Theory I B. Hayes

Class 10, 2/9/23: More on Phonotactics; Bases I

1. Current assignments

- Read Donca Steriade (1997) "Lexical conservatism and the notion base of affixation"
  - On web site
  - No summary required
- Current homework on phonotactics: **new deadline** Tues. Feb. 21.

MORE ON PHONOTACTICS IN GENERAL

2. This is an ancient topic

- B. L. Whorf published a formula for the English monosyllable in 1940 in *Technology Review*, the MIT alumni magazine:

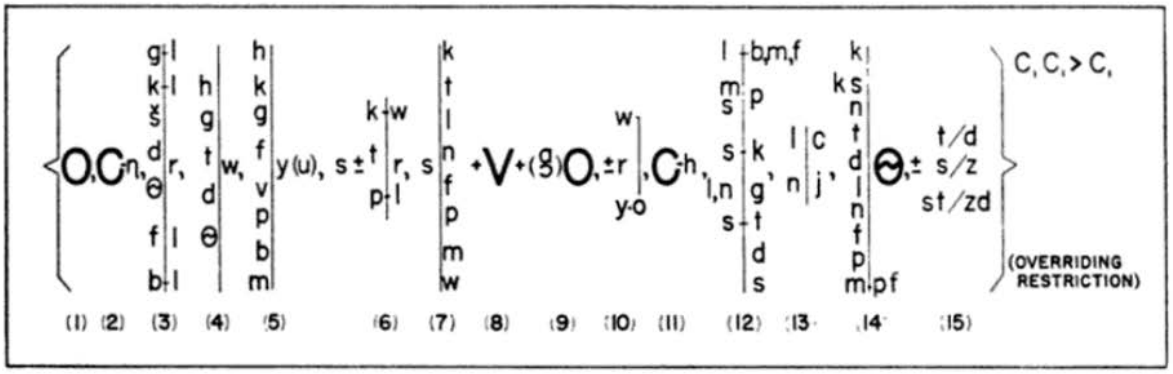

Figure 12. Structural formula of the monosyllabic word in English (standard mid-western American). The formula can be simplified by special symbols for certain groups of letters, but this simplification would make it harder to explain. The simplest possible formula for a monosyllabic word is $C + V$, and some languages actually conform to this. Polynesian has the next most simple formula, $O, C + V$.

Contrast this with the intricacy of English word structure, as shown above.

3. The Rich Base theory of phonotactics (Stampe 1970s, Prince and Smolensky 1993)

- The ranking of the universal constraint inventory defines a (non-probabilistic) filter through which the forms of the Rich Base must pass.
- This is a tough theory to probabilitize, since stochastic repair predicts free variation! /pɔɪk/ → [pɔɪk] or [paɪk] (or whatever)
- Our only hope of avoiding stochastic repair is giving up on analyzing intermediate well-formedness, which is supported by abundant experimental evidence.

4. The Evaluate-all-of-GEN theory of phonotactics

- I suspect the first to propose this was: Hayes, Bruce & Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39.379–440.
- Simply let the grammar assign a probability to every form in GEN.
- It is possible to use this to match lexical frequencies and even to some extent native speaker intuition.

5. A problem with this theory: Elsewhere

- Quechua has [e] as an allophone of /i/ adjacent to uvulars — it never occurs other than in this environment.
- Rich Base theory can get this: *qi > *e > IDENT(high) [☞ try the UR's /qi/, /qe/, and /e/, with high and nonhigh candidates].
- Evaluate-all-of-GEN must include *qi, *e, and a license: "FAVOR [qe]", missing a generalization. [☞ try the same candidates]
- For discussion see
  - Wilson, Colin, and Gillian Gallagher. "Accidental gaps and surface-based phonotactic learning: A case study of South Bolivian Quechua." Linguistic Inquiry 49, no. 3 (2018): 610-623.

6. How to employ maxent-over-GEN: two ways

- Difficult and thorough: the "UCLA Phonotactic Learner", written by Wilson for Hayes and Wilson (2008). Available at https://linguistics.ucla.edu/people/hayes/Phonotactics/index.htm
  - It uses a finite-state machine, following principles developed by Jason Eisner, to explore a vast space of whole-word candidates in feasible time.
- Simple and contingent: Find a subset of the phonotactics, hoping it is somewhat isolated from the rest of the system.
  - E.g., medial clusters, V … V sequences
  - With some patience, this can be done entirely with conventional office software.

BASES

7. Words are influenced by other words in their phonological form: cases

1. derivational morphology (base to derivative)
2. multiple bases in derivational morphology
3. derivation morphology dependent on the inflectional paradigm
4. systematic patterns of predictability in the inflectional paradigm

8. Words are influenced by other words in their phonological form: theories

- Theories for (1) above
  - *SPE* cyclicity and boundary symbols
  - Lexical Phonology

- Output-to-Output Correspondence (Benua 1995)
- Stratal OT (work of Kiparsky, Bermudez-Otero)

- Theories for (2) and (3):
  - Steriadean Lexical Conservatism, with modifications by Breiss
- Theories for (4):
  - Lexicalist inflectional morphology (the "morphome")

9. Warm-up exercise, from Leonard Bloomfield's *Language* (1933)

p. 366: "In the Central-Western type of American English, …"

[ă] is a plain allophone in some environments: / ___ rp, rk
dark, sharp
[ˈdăɪk], [ˈʃăɪp]

also, "before the clusters [rd, rt] followed by [non-suffixal] [-ɚ, -ṇ]"
barter, Carter, garden, marten (Martin)
[ˈbăɪrɚ], [ˈkăɪrɚ], [ˈgăɪdṇ], [ˈmăɪtṇ]

"Before a … suffix [-ɚ, -ṇ], however, the longer variant is used, as in
starter, carter ('one who carts'), *harden*
[ˈstaɪrɚ], [ˈkaɪrɚ] [ˈhaɪdṇ]

"Here the existence of the simple words *start, cart, hard* (whose [a] is not subject to shortening), has leave to the favoring of the normal, longer variant."

[ˈstaɪt], [ˈkaɪt] [ˈhaɪd]

- ☞ figure this one out with constraints and rankings. Be sure you rule out [ă] in monomorphemes.

10. A more carefully-documented case of this type

- Sugahara, M. & Turk, A. (2009) Durational correlates of English sublexical constituent structure. *Phonology*. 26, 03, p. 477-524

11. Warm-up exercise II: rendaku and [ŋ] in Conservative Tokyo Japanese

- Source: Junko Itô Armin Mester (1997) Correspondence and Compositionality: The Ga-gyo Variation in Japanese Phonology. In Roca, Iggy, ed. *Derivations and Constraints in Phonology.* Oxford University Press.

Rendaku (sequential voicing in compounds)

ta**m**a	'ball'	teppoo+**d**ama	'bullet'
ta**n**a	'shelf'	garasu+**d**ana	'glass shelf'

- For older speakers of Japanese, /g/ has a very interesting allophone in intervocalic position: [ŋ].
- This is typologically odd: most languages would lenite /g/ to [ɣ]. (is this moving the place of articulation closer?)

Distributional data, from Ito/Mester:

a. Initial g:

✓[**g** .....]	*[ŋ .....]	
geta	*ŋeta	'clogs'
giri	*ŋiri	'duty'
guchi	*ŋuchi	'complaint'
go	*ŋo	'(game of) Go'
garasu	*ŋarasu	'glass'

b. Internal ŋ:

*[... **g** ...]	✓[.... **ŋ** ....]	
*ka**g**i	ka**ŋ**i	'key'
*ka**g**o	ka**ŋ**o	'basket'
*ka**ŋg**ae	ka**ŋŋ**ae	'thought'
*sasa**g**eru	sasa**ŋ**eru	'give'
*u**g**uisu	u**ŋ**uisu	'(Japanese) bush warbler'
*toka**g**e	toka**ŋ**e	'lizard'
*i**g**irisu	i**ŋ**irisu	'England'

- Stem-final before suffix; suffix-initial:

oyo**ŋ**	+	oo	'swim-HORTATIVE'	stem-final
to**ŋ**	+	anai	'sharpen-NEG-PRESENT'	
kayoobi	+	**ŋ**a	'Tuesday-NOMINATIVE'	suffix/clitic-initial
mikka	+	**ŋ**urai	'approximately three days'	
gorira+no	+	**ŋ**otoshi	'like a gorilla'	

- Alternations in Sino-Japanese compounds, whose members are all *bound roots*:

Bound roots:

PrWd[**g**... .....]				PrWd[..... **ŋ** ..]			
gai	+	jiN	'foreigner'	koku	+	**ŋ**ai	'abroad'
go	+	zeN	'morning'	shoo	+	**ŋ**o	'noon'
gam	+	peki	'quay, jetty, wharf'	kai	+	**ŋ**aN	'sea shore'
gi	+	kai	'parliament'	shi**ŋ**	+	**ŋ**i	'deliberation'
guu	+	zeN	'accidental occurrence	soo	+	**ŋ**uu	'meet accidentally'
gen	+	zai	'currently'	sai	+	**ŋ**eN	'reappearance'

- How g ~ ŋ works in compounds whose members are **free stems**:

Compounding with g-initial $Stem_2$:		optional VVN			
geta	'clogs'	niwa	+ {**g** / **ŋ**}	eta	'garden clogs'
goro	'grounder'	pitchaa	+ {**g** / **ŋ**}	oro	'a grounder to the pitcher'
gara	'pattern'	shima	+ {**g** / **ŋ**}	ara	'striped pattern'
gei	'craft, art'	shirooto	+ {**g** / **ŋ**}	ei	'amateur's skill'
go	'Go game'	oki	+ {**g** / **ŋ**}	o	'Go played with a handicap'

- How k ~ ŋ works in compounds with a second-position k-stem:

Compounds involving Rendaku: obligatory VVN					
kuni	'country'	yuki	+ {**ŋ** / ***g**}	uni	'snow country'
kami	'paper'	ori	+ {**ŋ** / ***g**}	ami	'origami paper'
kaeru	'frog'	gama	+ {**ŋ** / ***g**}	aeru	'toad frog'
keŋka	'fight'	oyako	+ {**ŋ** / ***g**}	eŋka	'parent-child fights'
kaki	'writing'	yoko	+ {**ŋ** / ***g**}	aki	'horizontal writing'
kusuri	'medicine'	nuri	+ {**ŋ** / ***g**}	usuri	'medical ointment/cream'
kirai	'dislike'	onna	+ {**ŋ** / ***g**}	irai	'woman-hater, misogynist'

- Exercises:
  - formulate an OT analysis of these facts. I suggest *MAP constraints.
  - Reconstruct the historical chronology by which this pattern came to be.

- I suggest the following inputs, candidates, and constraints:

geta	**k**	
	g	**1**
	ŋ	
kagi	**k**	**1**
	g	
	ŋ	
niwa-geta	**k**	
	g	**1**
	ŋ	**1**
ori-kami	**k**	
	g	**1**
	ŋ	
ŋa (RB)	**k**	
	g	**1**
	ŋ	

*ng, *VgV, *Map-OO(g, ng), *Map-OO(k, ng), Rendaku
*Map-IO() for all three consonant pairs.

12. Ponder

Naturalness with respect to P-map

13. What have we learned?

- A paradigm uniformity effect — which is optional

TREATMENT OF PARADIGM UNIFORMITY EFFECTS IN *SPE* PHONOLOGY

14. The bifurcation

- Inheritance of derived phonological properties: the **cycle**
- Resistance to acquisition of properties: **word-internal boundaries**.

15. Cyclic effects

- English secondary stresses are (roughly) left-to-right binary, no clash, in the pretonic domain.

Examples from Hayes (1982, *LI*).

àbracadábra	Kàlamazóo
Lùxipalílla	Hàrdecanúte
Pèmigewássett	Àllamakée
Òkefenókee	Ìllilouétte
Nèbuchadnézzar	Màttamuskéet
pàraphernália	Àntigonísh
Kìlimanjáro	Gàllipolís

Pòpocàtepétl	Òkalòacóochee	Àpalàchicóla	Àntanànarívo
Hànamànióa	ìpecàcuána	ònomàtopóeia	hàmamèlidánthemum

This is not respected in suffixed forms, where there is a cyclic effect present:

*sublìmináli*ty
*demòcratizá*tion
*Macàssaré*se

- What the analysis has to do: Faithfulness of *sublìmináli*ty to *sublíminal* has to outrank the rhythmic principles that determine in monomorphemes.
- *SPE*'s method: cyclicity, on bracketed structure (brief demo)

16. Boundary effects: the distribution of preantepenultimate stress

- There are no stems whatever ending in stressed plus three stressless: "Hi, I'm *[ˈpæmələnə]"
- With **productive suffixes**, pre-antepenultimate stress seems rather normal and possible in new words:

-ing	*monitoring, jettisoning*
-eth	*seventieth*
-ish	*Madison-ish*

17. The *SPE* analysis of boundary effects

- *P*roductive suffixes are treated with "#". "Readjustment rules" apply.
  - Rule 1: [ ] → [# #]
  - Rule set 2: X #] ation → X] ation; etc., for the less-productive affixes.
  - Stress rules apply in domains bounded by # #.
  - In translated form (prosodic structure), this is still a living analytic option, see e.g. Peperkamp, S. (1997). *Prosodic Words*. HIL dissertations 34. The Hague: Holland Academic Graphics.

18. Paradigm uniformity effects not covered in *SPE*

- These words seems to have influence from their base forms.

- But there are funny relations to the base, e.g. truncation of affixes, semantically inappropriate base
  - We will study Steriade's views on these later on.
- Below is a sorted list from a dictionary search for preantepenultimate stress words

19. -able forms with Preantepenultimate Stress

abominable
applicable
communicable
estimable
inalienable
incalculable
inextricable
inseparable
interminable
inviolable
navigable

How to form?

20. Some -ative form

communicative
palliative
speculative
cumulative

21. Some -acy forms

What is the pattern here?

accuracy
adequacy
advocacy
candidacy
celibacy
confederacy
degeneracy
delicacy
immediacy
intimacy
intricacy
legitimacy
literacy
obstinacy

22. Upshot

- We are seeing not the straightforward inside-to-outside derivations proposed in *SPE*.
- Rather, various quirky relationships within the derivational paradigm — which we might plausibly be able to get with specific OO-correspondences constraints.

23. The apparent virtue of the Paradigm Uniformity approach

- It unifies:
  - affix neutrality (*SPE* boundaries)
  - simple inside-to-outside influences (*SPE* cyclicity)
  - quirky correspondence relations (*SPE* ignored these)

24. Rich variety of Paradigm Uniformity types

- Paradigm Uniformity is sensitive to the paradigm involved; i.e. we may need to be quite specific about the morphological relations present. E.g.
  - -*ing* is totally straightforward in selected a sensible, local base and maintaining its phonology: *ímitàting*
  - -*able* is sometimes -*ing*-like, but can also use -*ate*-truncated bases. It never changes base stress, as far as I know.
  - -*ian*, used by academics, is quite productive and can even resurrect lost vowel qualities: *Abelian*, *Sokolian*, *Kruskalian*

- This is perhaps a theme in the use of OO correspondence — it often seems to be process-specific.

25. Returning to the Japanese example

- It's worth thinking about Japanese from the viewpoint of *SPE*, for it involves something they don't/can't cover.
- "Optional cyclicity": a rule must be made optional when it applies on an inner cycle, but then obligatory when the stem occurs by itself. [☞ go through this]
- This seems at least inelegant to me: what is actually optional is whether you carry forward the effects of the base form on the derived form, per the principle of constrint ranking.
- This critique holds not just of *SPE*, but of Stratal OT, which retains the cyclicity approach to paradigm uniformity.