Linguistics 219 Spring 2020
Phonological Theory III B. Hayes

Class 19, 6/2/2020: Gradient Symbolic Representations

1. Assignments

- Read:
  - Matt Goldrick and Paul Smolensky (2016.) Gradient symbolic representations in phonology: The case of French liaison. ROA 1552.
  - On course web site.
- Feel free to discuss progress/problems with your term project.
  - Wednesday office hours at regular hour.
  - **Friday office hours moved to 4, not 2.**
- Handout-session with me before the middle of Finals Week, Wed. 6/10.

THE WORK

2. The idea

- Throughout this course we have worked with **grammars** that are gradient (probabilistic).
- Here, the focus is on **representations** that are gradient: "this position has in it something that has a probability of .3 of being an [n], .3 of being a [t], .3 of being a [z]"

3. Paul Smolensky

- In addition to being a co-inventor of OT, he is an acclaimed cognitive scientist (Rumelhart Award winner, etc.).
- His cognitive work, since 1986, has attempted to create an intelligible connectionism
  - See, notably, his massive book with Legendre (2006) *The Harmonic Mind*
- He is arguably the inventor of MaxEnt, in pre-linguistic work from 1986
  - Smolensky, P. "Information processing in dynamical systems: foundations of harmony theory." In *Parallel distributed processing: explorations in the microstructure of cognition, vol. 1: foundations*, pp. 194-281. 1986. which I encourage students to cite for MaxEnt
  - but seems to have disavowed it, at least as far as the present paper is concerned

4. Status of the Goldrick/Smolensky paper

- The work is cited to a fair degree and likely to be at least mentioned at any phonology conference.
- However, it remains a ROA manuscript, for reasons unknown to me.

WHAT IS FRENCH LIAISON?

5. A bit on data

- I found it useful to compare G+S's presentation with the data section of
  - (2017) Kie Zuraw and Bruce Hayes, Intersecting constraint families: an argument for Harmonic Grammar. Language 93: 497-548.

  (written entirely by Kie).
- I have pasted some examples gathered by Kie to illustrate the same points made by G+S.

6. Historical change dishes up wug tests flunked by children

- We have seen this in various ways in this course — a valuable source of ecologically valid natural experiments.
- We have also see the mechanism creating this: postlexical change applies to some degree blind to its consequences.

7. The evident historical origin of Liaison

- French is a Romance language whose history shows heavy erosion at the right edges of words, of both vowels and consonants.
- It also shares the Romance characteristic of syllabifying across word boundaries.
  - Harris (1983)'s lovely Spanish sentence

    lo.s o.tro.s es.ta.βa.n e.n e.l a.βjon
    'the others were on the airplane'

- This transword syllabification *bled* the deletion of word-final coda consonants, creating alternations.

'little-m.'	'little-m friend-m'	
pətit	pətit ami	ancient original forms
pə.tit	pə.ti.t a.mi	shown with syllabification
pə.ti	—	Loss of word final coda consonants

8. Liaison is sometime very lexical

- Feminine article is [l] before vowel, [la] before consonant.
- "of the-masc.' is [dy] before consonant, [də l] before vowel.
- etc., primarily for grammatical words

9. Opacation I: the loss of final schwa

- Think of the Romance theme vowels: *-o* and *-a*, as in Spanish/Italian.
- Before any of this stuff happened, the *-o* fell completely.

- *-a*, more sonorous, survived to a later date in the form of reduced -ə.[1]
- This -ə survives today very marginally — e.g. in poetry and song — but basically fell — *after* the loss of final C.

'little-feminine'	'little-masculine'	
pə.ti.t-ə	pə.tit	early forms syllabified
—	pe.ti	Final Consonant Drop
pə.tit	—	Final Schwa Drop

- Thus, the loss of final consonants is opacated.
  - ➢ [☞ Classify this one in classical Kiparskian terms.]

10. Opacation II: the loss of initial [h]: *"h aspiré"*

- Initial [h] fell, again after Final Consonant Drop

'little'	'little friend'	'little hero'	
pə.tit	pə.tit ko.pɛ̃	pə.tit. he.ro	Ur-forms, syllabified
pə.ti	pə.ti ko.pɛ̃	pə.ti. he.ro	Final Consonant Drop
—	—	pə.ti e.ro	h Drop

[☞ Classify this one in classical Kiparskian terms.]

11. Opacation III: creation of new consonants by Glide Formation

i u y → j w ɥ before vowel
These glides contrast with the glides that were always glides.

le yodle	[**lə** jɔdl]	'yodels it'	**l'**iode	[**l** jɔd]	'the iodine'
le yaourt	[**lə** jauʁt]	'the yogurt'	**l'**yeuse	[**l** jøz]	'the oak'
la hiérarchie	[**la** jeʁaʁʃi]	'the hierarchy'	**l'**hiatus	[**l** jatys]	'the hiatus'
la huée	[**la** ɥe]	'the booing'	**l'**huître	[**l** ɥitʁ]	'the oyster'
le huitième	[**lə** ɥitjɛm]	'the eighth'	**l'**huile	[**l** ɥil]	'the oil'
			l'huissier	[**l** ɥisje]	'the bailiff'
le ouistiti	[**lə** wistiti]	'the marmoset'	**l'**ouest	[**l** wɛst]	'the west'

[1] French schwa is a short, reduced vowel, but with phonetic identity [œ]. I follow tradition in calling it schwa.

12. Responses of French-learning children to opacity I: limitation of the inventory of consonants that alternate by Liaison

- Originally, it seems, quite a few consonants alternated with zero.
- There are a few lexically-listed relics that still alternate other consonants, to some extent, at least in normative speech.

[p]	trop ~ tro	'too much'	*trop*
[g]	lɔ̃g ~ lɔ̃	'long'	*long*

- But now, only three consonants alternate.

[n]	contextual: after nasal vowels	mɔ̃n ~ mɔ̃	'my'
[z]	very often morphological; e.g. in plurals		
[t]	perhaps the true default		

- Cf. Polynesian languages, which likewise lose final consonants, preserving them before vowel-initial affixes.
  - The set of surviving alternations shrank, approaching one single (language-specific) alternator (Hale 1973).

13. Responses of the French kids to opacity II: the partial decoupling of h-aspiré status from etymological initial [h]

- New words sometime enter the language as h-aspiré even though they don't begin orthographically with an *h*.

la hache	[**la** aʃ]	'the axe'
du haricot	[**dy** aʁiko]	'of the bean'
le uhlan	[**lə** ylã]	'the uhlan (central/east European lancer)'

- Ditto for the glide cases above.

14. A (more or less) *non*-response to opacity: no innovating epenthesis

- ʒo.li a.mi 'good-looking friend' is preserved.
  *ʒoli t ami

- Compare restructured [r] epenthesis in nonrhotic English varieties; [l] epenthesis in some Midwestern varieties that have lost final [l].

15. An interfering factor: French culture and education

- Opacity repair is clearly happening, but is retarded by the world's strongest tradition of normative grammar.

- Archaic forms of liaison are counted as "cultivated" and "correct" for French speakers.
- Unguarded vernacular speech displays the innovations.
- Were French a vernacular-only language, we might imagine that repair might have happened much faster.

THE INTRIGUING PHENOMENA (G+S)

16. Retention of the final consonant phrase-finally in a few words

ɥit 'eight'
ɥi.t a.mi 'eight friends' (I think; I provided this form.)
ɥi. ko.pɛ̃ 'eight pals' (I think; I provided this form.)

Three other words, unspecified, are said to apply this way.

Conclusion: Liaison is a property of Word1.

17. Presidents of France are allowed to liase-but-not-resyllabify

ʒ-a.ve.z … ɛ̃ rɛv
I had a dream

Conclusion: Liaison is a property of Word1.

18. Similarly: pauses after the liaison consonant

sɑ̃**z** … ɑ̃vizaʒ-e
without expect-inf. 'without expecting'

Conclusion: Liaison is a property of Word1.

19. Similarly: pauses before the liaison consonant

pətit- … z istwaːʁ
little-pl. story-null plural 'little stories'

Conclusion: Liaison is a property of Word2.

- I wish they would tell us the circumstances under which you get (18) or (19).

20. Frequent phrases

- If Word1 Word2 form a commonly uttered phras e, then Liaison is far more likely.

Many authors (e.g. L'Huillier 1999:37): *huit* 'eight' behaves as unaspirated only within certai numerals (e.g. *dix-huit* 'eighteen' (liaison), *vingt-huit* 'twenty-eight' (liaison), *trente-huit* ' (schwa deletion)); but as aspirated, with no liaison, in *cent huit* 'one hundred and eight'.

PERFORMANCE-Y PATTERNS

21. Errors by high school students when tested

In one class of error, in a location where a given /L/ consonant should appear, an incorrect consonant appears in its place. In such errors, the erroneous consonant is almost always another liaison consonant. Quite a few errors of this type occurred in a reading study of French high-school students reported by David Hornsby (2011) [25]; e.g., the sequence *long apprentissage* 'long apprenticeship' was pronounced correctly (with /L/ = [g]) 4 times and incorrectly 20 times — all errors of this type. Across the entire experiment there were 81 errors of incorrect C insertion, and the erroneous consonants were: [z] (56 times), [t] (23 times), or [n] (2 times) — in every case, one of the 3 productive liaison consonants.

22. Contra above, children *do* insert consonants in etymologically hiatic environments

as in ʒoli ami 'cute friend', with t z n

23. Children use a Liaison consonant postpausally

tami, zami, nami

- Such "recutting" has happened a number of times in English: English *nuncle* from *mine uncle*, *thine uncle*).

Conclusion: Liaison is a property of Word2, or at least kids entertain this as a possibility.

ANALYSIS

24. Target derivation

- What is being derived is *petit ami* [pə.ti. t-a.mi] 'little friend-masc.'

25. *petit* ends in a gradient consonant

- p. 15 gives a sketch of three kinds of word.

peti(t)	/pøti(λ·t)/
juste	/ʒys(1·t)/
joli	/ʒoli/

- *petit* has a final semi-t with an **activation value**[2] of λ
  - λ is later assumed to be .5

- For *juste*, with invariant [t], λ is 1.
- For *joli*, with no [t] at all, there is no consonant represented.
  - I guess a [t] with λ = 0 would be equivalent?

26. *ami* begins in a gradient consonant

ami	/ℒami/	where ℒ ≡ (τ·t + ζ·z + ν·n)
héro	/eʁo/	(*h-aspiré*)
copain	/kopɛ̃/	

- Where:
  - tau is the activation for beginning with a [t]
  - zeta is the activation for beginning with a [z]
  - nu is the activation for beginning with an [n]
- These values are set at .3.

- h-aspiré words are represented simply as themselves, without any hidden semi-consonants.
- Consonant-initial words are represented simply as themselves.

27. Liaison is the merger of two abstract entities

- Merger is common currency in phonology, dating to Panini.
- E.g. French acquired its [o] from earlier [au], still reflected in the spelling (e.g. *Laurent* [loʁɑ̃])
- Tagalog famously merges /ŋ/ with following consonants: /ŋp/ → [m].
- Etc.

28. The merger needed for Liaison

A liaison consonant $\mathcal{L}$ that surfaces in the sequence $W_1W_2 = \hat{W}_1\mathcal{L}\hat{W}_2$ derives simultaneously from two underlyingly sources: $/W_1/ = /\hat{W}_1(\lambda\cdot\mathcal{L})/$ and $/W_2/ = /(\gamma\cdot\mathcal{L})\hat{W}_2/$, where W_1's underlying weak final $\mathcal{L}$ and W_2's underlying weak initial $\mathcal{L}$ have activity values λ and γ respectively.

[When $\mathcal{L}$ = t, we write γ as 'τ'; when $\mathcal{L}$ = z, as 'ζ'; and when $\mathcal{L}$ = n, as 'ν'.]

- We see again the activation values
  - λ, for "semipresent" final consonants.

[2] A widely used term (idea?) in psychology. Activation can be given specific meanings in particular models, such as neural network models, or G+S's model.

- tau, for "semipresent" initial [t]
- zeta, for "semipresent" initial [z]
- nu, for "semipresent" initial [n]

29. A tableau

Gradient Harmonic Grammar tableau for configuration vCv (*petit ami*)

	weight: [ᴹpøti(λ·t$_1$)] [ᴹ(τ·t$_2$+ζ·z$_3$+ν·n$_4$)ami]	*–10* DEP	*2* MAX	*1* ALIGN-L	*–0.9* ONSET	*–0.7* UNIF	*H*
a	.pø.ti.a.mi.				1		–0.9
b	.pø.ti.t$_{12}$a.mi. ☞	1–(λ+τ) 0.2	λ+τ 0.8	1		1	**–0.1**
c	.pø.ti.t$_1$a.mi.	1–λ 0.5	λ 0.5				–4

- Let us take this apart to understand it fully.

30. The constraint inventory

- In name at least, these are standard-issue, adopting the classical McCarthy-Prince Faithfulness theory.
- But **evaluating their violations** in the new theory involves us in novel issues; see below.

31. The weights

- G+S adopt the widespread (non-Wilsonian) convention that negative values are penalties, total harmony is normally negative, and that the highest value of Harmony is the winner.
- Let's check with Candidate 2:
  - 0.2 (we'll cover this) x −10 = −2
  - 0.8 x 2 = 1.6
  - 1 x 1 = 1
  - −0.7 x 1 = −0.7
  - Total = −0.1, as promised

32. The input

- This is formed by concatenating *petit*, complete with its gradient stuff …
- with *ami*, complete with *its* gradient stuff

33. The candidates

- *[pøti ami], dropping all the gradient stuff
- ☞ [pøti t$_{12}$ ami], merging the gradient stuff into an authentic segment [t].
- *[pøti t$_1$ ami] attempting to let just the gradient stuff of *petit* get to the surface on its own.

- I.e., it would be homophonous with the winner, but the grammar doesn't (and shouldn't) generate it.
- Presumably *[pøti t_2 ami], *[pøti z_2 ami], *[pøti n_2 ami] would get essentially the same analysis.

- I believe all candidates must be **pronounceable**; i.e. activations must be one.

34. The violations: easy cases

- ONSET: patently violated by [pəti ami].
- UNIF: this is UNIFORMITY, from McCarthy and Prince, banning output segments with two input correspondents.

35. Redefining MAX

- MAX is a **positive constraint**, seen elsewhere in this course (e.g. Flemming, your own CC charts).
- "MAX rewards underlying activity that makes it to the surface."
- As you can see, you get Max "violations" by summing up the lamba/tau/zeta/nu activation values.
- We might assume that we also sum up the 1's of non-fancy segments, but these will cancel out across candidates, so this can be skipped.

36. ALIGN is also a rewarder

- ALIGN is related to classical ALIGN(Morpheme, L, Syllable, L): "Penalize every left-edge of a morpheme that fails to coincide with the left edge of a syllable"
  - Yet it too is recast as a reward.
  - "… rewards a morpheme if the output correspondent of a segment at its left edge falls at the left edge of a syllable"
  - the surface [t] of winner [pətitami] comes, in part, from the first segment of ami (the gradient segment), and, unlike in the other two candidates, it obtains a reward for surfacing at the beginning of a syllable.
- All "violations" common to the three candidates are harmlessly ignored here.
- Is there any work in the literature exploring the general consequences of recasting ALIGN constraints as rewarders?

37. Redefining DEP gradiently

- Analogy: when you insert a segment entirely, you incur a violation of 1 for DEP.
- If the segment is "partly there" in the UR, then you incur a *lesser* violation of DEP:
  - One minus the underlying activation of the segment (= how far you have to go to 'get there').
  - Further: for merged segments, we sum the activation of the underlying segments, and subtract this from one to get the DEP violation count.

38. Defining winners

- This is **non-stochastic** Harmonic Grammar: you need only achieve better Harmony than any other candidate to win.
- So I think we now have the whole thing. Take another look.

39. Exercise I: evaluate *[pøtit$_1$ t$_2$ami]

40. Exercise II: is [peti t2 ami] harmonically bounded?

- What are the consequences of being harmonically bounded?

41. Exercise III: is [peti ami] harmonically bounded?

42. Exercise IV: what happens to *petit* in prepausal position (alone)?

43. Exercise VI: what is the mechanism for blocking Liaison before an h-aspiré word like *heros*?

PERFORMANCE ISSUES

44. Here, gradience returns

- The familiar-to-us formula $e^{H(r)}$ is introduced, which with a Z term will yield probability.
- Goldrick is coauthor of a quite interesting paper that uses MaxEnt to model speech errors:
  - Goldrick, Matthew, and Robert Daland. "Linking speech errors and phonological grammars: Insights from Harmonic Grammar networks." *Phonology* 26, no. 1 (2009): 147-185.
- So now, as familiar to us, candidates that have pretty-good harmony will get pretty-good probability, and be likely as speech errors.

COMMENTS ON THE WORK: GRADIENCE

45. In the title, "gradient" has a completely different meaning from the rest of this course

- Theoretical framework is Classical Harmonic Grammar, with a single winner.
- For cases of variation in language (e.g. variable outputs, gradient well-formedness, lexical frequency matching?) we are recommended to adopt the (never-implemented) Multiple Grammars approach.
- The gradience of representations is solely **diacritic**; it distinguishes a variety of distinct behaviors.
  - It might have been clearer to restrict the lexical activation values to integers.
  - See also G+S's mention of Zuraw, and Hayes (2017), where our purely-for-convenience division of the gradient data into five groups is taken as a serious proposal.

46. Uses of gradience in G+S

- Speech errors, as above
- Acquisition

47. A sketch gradient acquisition algorithm (G+S)

- The child must eventually learn that:
  - [ami] begins with the hidden-structure tripleton.
  - [ero] begins with nothing.
- Schematic picture:
  - First Chouchou hears [pøti] and [pøtitami], yielding lexical representations /pøti/ and /tami/ (the latter is indeed a possible child production error, per above).
  - Then she hears [ʒoli ami]. Help! Where's the [t]?
  - So she weakens the /t/ in the lexical representation of /ami/, making it gradient.
  - Presumably [lez ami] 'the friends' and [mɔ̃ nami] strengthen the gradient representation of [z] and [n]/
- This idea is acknowledged to be schematic — learning hidden structure is hard, and surely they've got hidden structure in spades.

48. The above is a tadpole/frog theory

- Metaphor from work of Lila Gleitman in the early 1980's.
- I.e. acquisition of gradient representations follows a normal probabilistic path.
- Then, somehow, at the end, the variation that the child encounters is re-sorted into a multiple-rival-grammars framework.
- I would enjoy seeing this scheme worked out and supported empirically.

49. What belongs in the lexicon, what in the grammar?

- A powerful generalization, extending itself and found in productive errors (above), is:
  - Liaison only for [t], [z], [n].
- The theory expresses this principle as a separate installation, for hundreds of words, of the special abstract initial semiconsonant with activation .3 for [t], [z], and [n].
- This is perhaps not wrong, but it is certainly **nonstandard** in the historical context of phonological analysis:
  - Across the board generalizations go in the grammar, not in the lexicon.
  - See Kenstowicz and Kisseberth (1977, 1979 books) for clear statements of such principles.
- Note that the apparatus for installing the key generalization *in the grammar* already exists, e.g.
  - *MAP(p ~ ∅), *MAP(d ~ ∅),*MAP(l ~ ∅), etc. are undominated.
  - *MAP(t ~ ∅), *MAP(z ~ ∅), *MAP(n ~ ∅) are weighted low enough to be violated in winners.

- Perhaps the crucial generalization holds over the **segment inventory**?
  - "The consonants of French are: [p t k b d g f s ʃ v z ʒ m n ɲ l ʁ w j, λ x t, tau x t, zeta x z, nu times n]
  - But then we have to say something about the Rich Base, perhaps having a Stem Level that only permits the needed consonants (including the fancy abstract ones) to survive.

50. What is the general applicability of this theory?

- We are given a clear hint: phenomena with two competing, distinct explanations.
- Here:
  - The liaison consonant behaves like part of Word1.
  - The liaison consonant behaves like part of Word2.
- Elsewhere, Dowty (2003) is cited as a model.
  - PP's that behave sort of like an argument and sort of like an adjunct.

51. Relating to some local work

- Lexical propensity (Zuraw and Hayes 2017, Zymet dissertation 2018)
- The possibility that every word has a particular propensity to induce some particular alternation.
- This might make the word-specific numbers like lamba, tau, zeta, nu look more appealing, contra the above.
- Nevertheless, the authors just cited work this by giving words specific weights for specific constraints in the grammar; not activiations for individual segments.
- I suggest this may be the best way to achieve the needed balance between lexical idiosyncracy and pan-grammar behavior.