

## Class 7, 11/13/20: Last stress items; Experimental work

### 1. Assignments

- Readings:
  - Read Kelly, Michael. 2004. Word onset patterns and lexical stress in English. *Journal of Memory and Language* 50:231-244.
- Homework #2, due today
- Term paper —talk with me about your topic if you haven't yet.

### 2. Readings so far

- SPE Chap. 2
- Carlos Gussenhoven (1982) English plosive allophones and ambisyllabicity. *Gramma* 10:119-141. Download below.  
Timothy Vance (1987) 'Canadian Raising' in some dialects of the northern United States. *American Speech* 63:195-210.
- Liberman and Prince (1977, *LI*) "On stress and linguistic rhythm", up to p. 309.
- Pater, Joe (2000) Non-uniformity in English secondary stress : the role of ranked and lexically specific constraints. *Phonology* 17:237-274.
- Moore-Cantwell, Claire (2016) The Representation of Probabilistic Phonological Patterns: Neurological, Behavioral, and Computational Evidence from the English Stress System. Read chs. 3 and 4.

### 3. Looking ahead

- Segmental phonology, such as Vowel Shift and Velar Softening
  - Experiments, such as Pierrehumbert
- The past tense system, and trying to glean something useful out of the intellectual wars surrounding it.

## FILLING IN SOME GAPS; PATER-REPAIR

### MORE ON THE EFFECT OF SYLLABIC SONORANTS

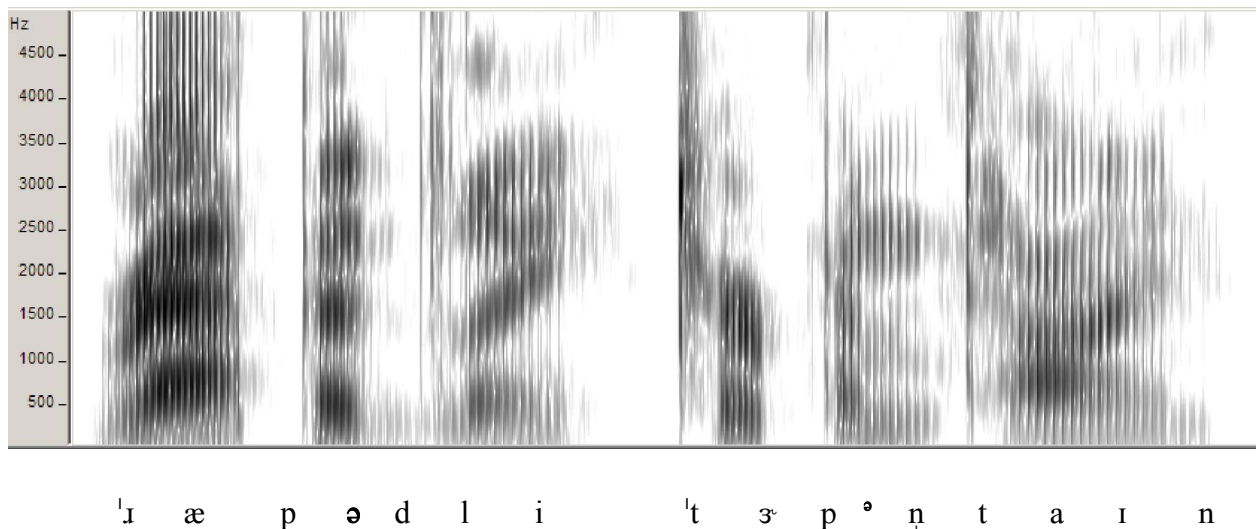
### 4. What Pater has accomplished — summary

- He covers the *pretonic* regions, focusing on the contrast exemplified by:
  - <sub>1</sub>Fran'cisco vs. <sub>1</sub>San Fran'cisco
- He has a beautiful idea for why this would occur specifically in sonorant-closed syllables.

- On the face of it, *contra naturam*, since in some languages (Kwakwala), sonorant codas make their syllable heavier than obstruent codas.
- The key is to treat the sonorant as surface as (essentially) syllabic, hence light.
- So then the grammar becomes a decision about when you want to opt for this.
- You do it preferentially in *San Francisco* because it doesn't create an extra unfooted syllable, as it would in *Francisco*.

## 5. Is the relevant sonorant truly syllabic?

- Indubitable cases:
  - r: *conversation* əˈvɜːʃən
  - l, I think: *cavalcade* [l]ˈkæ.vəlˌkeɪd
  - n, after an alveolar: *potentate* [n]ˈpɒ.tənˌteɪt
- A less likely case: *turpentine*, here compared with *rapidly*



## 6. Kiparsky's (1989) discovery about Gerard Manley Hopkins (1844-1889)

- He is, unusually, a syllable-weight poet working in English.
- His criterion for weight is special for stressless syllables that end in sonorants — intermediate!
- thus, *dapple* is either [dæ.pəl], light-heavy; or [dæp˩l], light-light.
- For details and corpus search, see Hayes and Moore-Cantwell's MaxEntification of Kiparsky, in *Phonology* (2011)

## 7. Extending Pater, based on earlier literature

- I believe that the *Francisco/San Francisco* distinction holds for *primary* as well as secondary stress (see Hayes 1982, *LI*, where this is laid out).
- Let us set up the scenario:

1. There is final secondary stress, either due to a long vowel, or Russian constraints (clusters, non-coronal, lexical listing)
  2. The penult is CVC, closed by a sonorant.
  3. There is just one remaining syllable to the left.
- If you have all three, then the light syllabic sonorant option is normally chosen, and you get antepenultimate stress.

## 8. Lexical search

- Sonorant-closed penult, final (or penultimate, with *-ary* etc.) secondary stress, trisyllabic
  - Stress should be penultimate, with the Paterian configuration
  - True for 37 words in the database:

*argentine, cavalcade, celandine, circumcise, columbine, commentary, commentator, compensate, davenport, desultory, dysentery, edelweiss, Fahrenheit, fragmentary, frankincense, inculcate, inundate, involuntary, legendary, mackintosh, melancholy, mercantile, merchandise, momentary, nightingale, nincompoop, potentate, quarantine, recompense, reconcile, recondite, secondary, sedentary, serpentine, turpentine, valentine, voluntary*

- False for just 2 words, both plausibly inheritance cases: *enfranchise, percentile*
- Same setup, but obstruent-closed penult. Penult should be stressed., single syllable precedes: should be skipped
  - True for 9 words: *monoxide, decathlon, Aquidneck, Penobscot, Hopatcong, Monadnock, jujitsu, Zimbabwe, Ojibway*
  - False for 4 words: *anecdote, designate, recognize, satisfy*
  - But we have independent reason to think that *-ate* and *-ize* are antepenultimate stress assigners.
  - Perhaps the fact that the initial syllable is light (“Arab” Rule) also matters.
- Same setup, but one more syllable on the left, analogous to Pater’s *ˌHaliˌcarˈnassus*
  - *Agaˈmemnon, Eniˈweˌtok, ˌMassaˈpeˌquod, ˌMemphreˈmaˌgog*

## 9. What is the productive pattern? Blick-test yourself

- Order of cases given is, sonorant-postlight, sonorant-postheavy, obstruent post-light, obstruent post-heavy.

### Should be antepenult

*Heppalcong, Hompalcong,  
Abelmosk, Ambelmosk  
Embendome*

### Should be penultimate

*Hempatcong  
ambecmosk  
Embecdome*

*Es'cumi'nac**,Escu'men,dac***10. What about modeling?**

- Pater's system, without modification, derives the wrong answer for the crucial (because reasonably frequent) trisyllabic forms.
- I found that it's not hard to fix: we split \*CLASHHEAD into two versions:
  - Strong version for when the clash is *after* the head;
  - Weak version for when the clash is *before* the head:
- Tableaux:

/palindrome/	Stressless VV	*Post Clash Head	Ident-IO	Non-Fin	Align-Head	Parse \$	Weight To Stress	*Pre Clash Head	*SonNuc	Align-L
☞ [1palN][2drome]				*	**				*	*
[1palin][2drome]				*	**		!			*
pa[1lin][2drome]		*!		*	*	*				**
pa[1lin]drome	*!				*	**	**			**

/agenda/	Stressless VV	*Post Clash Head	Ident-IO	Non-Fin	Align-Head	Parse \$	Weight To Stress	*Pre Clash Head	*SonNuc	Align-L
☞ a[1gen]da					*					*
[1agN]da					**				*	
a[1genda]				*!	!					*

### 11. What about modeling II? the MaxEnt model being developed here

- Hard to implement thoroughly, since we don't know when a final syllable is stressed except when it has a long vowel.
- Considering just the latter cases, a constraint that says:

“Don't have penultimate stress in a trisyllable not closed by an obstruent”

- This gets a decent weight; and behaves correctly:

*pendenmote*    favors antepenultimate stress  
*pendenma*     favors penultimate stress (but not by enough)

### 12. The wise pioneers again

- Once again, we have a pattern where the lexicon does quite a poor job of training the learner, and we wonder if (assuming productivity), the learner manages to pick the pattern from more subtle cues.

### 13. What is remarkable to me about this

- Presence of secondary stress on the final syllable determines where primary stress goes.

## UPGRADING THE MAXENT MODEL IN PROGRESS

### 14. Some notes on what it is allowed to refer to

- Segmentism and (plausible, but guessed) syllabification.
- Part of speech
- Presence of particular affixes, insofar as orthographic strings usually suffice to identify them.

### 15. What is excluded as cheating

- Use of orthography as clue to stress, as in *an'tenna* vs. *'Helena*.
- Obviously, use of the stress marks in the transcription.<sup>1</sup>
- “Inverse” use of Vowel Reduction as a clue to stress: AH is likely to be a schwa (AH0), not a stressed caret (AH1, AH2).
  - It is helpful that the CMUbet uses the same symbol for both.

---

<sup>1</sup> I am tempted to try a limited use, namely secondary stress on final syllables, for Sonorant Retraction (see below).

## 16. The GEN function

- This is ad hoc and frankly explorative:
  - Only Main stress is varied ( $n$  candidates for words of  $n$  syllables).
  - For the true way to do it, see Moore-Cantwell readings: every combination of 1, 2, and 0, with maximum one main stress.
- Indeed, this essential even in a *main stress* model, for there is one case (just given) in which the position of the main stress is strongly influenced by the position of the secondary stress.

## 17. This grammar is as far as we got last week

Constraint	Wght	Viol. Count	Description	Violating forms
*SkipHeavy	2.16	1649	Don't have a heavy syllable in the middle of a trisyllabic LP foot (SPE)	*galaxy, *colander
NonFin	2.53	1681	No final main stress	*balloon, *Tennessee
NonFinPoly	2.22	177	Same but for longer words only; see below (me ??)	*Tennessee
AlignR-3	-1.48	5007	"Penalizes" (rewards) antepenultimate stress, confirming <i>SPE</i> and other work	*cinema
AlignR-4	1.06	671	"Penalizes" (rewards) preantepenultimate stress (Pater 2000)	*hesitancy
AlignR-5	1.67	41	ditto for 5-ult (Pater 2000)	*amphitheatre, cannibalism, caricaturist, circulatory
Suf -ic	4.72	13	Penalizes any word with the orthographic suffix -ic if stress is not penultimate (SPE)	*Arabic
VerbFin	1.65	1015	Penalize verbs that don't have final stress (SPE)	*banish
SupHFin	2.26	2443	Penalize words that end in VV(C) or CC that don't have final stress (SPE)	*interest
Conjoin	0.52	521	Conjoins the last two constraints	*capsize, *govern
Hiatus	3.85	39	Don't stress an [i] or [ə] that is in hiatus (Karvonen, for Finnish)	*Korea, occurrence
CM-C[i]	1.89	70	Don't stress the penult in a trisyllabic or longer word that ends in [i] (Claire Moore-Cantwell)	*apache, bikini, canary
SonDestr	1.31	65	If a word is trisyllabic, has a VV	*incarnate, older

			final syllable and a VR medial syllable, it must have antepenultimate stress (Hayes 1982)	*infiltrate
PalAlv	4.19	22	If the final syllable begins with [ʃ] or [ʒ], stress must be penultimate (me)	*protege

## 18. Review: making weights intuitive

- The method is:
  - Two candidates with comparable violations
  - but one violates an extra constraint with weight  $w$
  - take  $e$  to the  $w$
  - That is the **odds** of the two candidates.
- Here are some weights converted to odds:

0.1	1.11
0.2	1.22
0.3	1.35
0.4	1.49
0.5	1.65
0.6	1.82
0.7	2.01
0.8	2.23
0.9	2.46
1	2.72
2	7.39
3	20.09
4	54.60
5	148.41
6	403.43
7	1096.63
8	2980.96

## A NOTE ON MODEL EVOLUTION

### 19. An example of the effects of refining the model

- Premise: the system prefers antepenultimate stress when all else is equal.
- Suppose we have *not yet* dealt with prestressing suffixes like *-ic*.
- What is the effect of having, or not having, this constraint.

## 20. A method adopted from Moore-Cantwell (readings, p. 139)

- What is the *generic* outcome, calculable by excluding all irrelevant factors.
- Here, we can try a form like /pæmɛlɒnʌ/, which the most important applicable constraints are:

<b>-1.38</b>	<b>2.09</b>	<b>1.92</b>
AlignR-3	AlignR-4	AlignR-5

- For this word, we will get this probability distribution for the four possible stressings:

/pæ	mɛ	lɒ	nʌ/
0.058	0.752	0.190	0.001

(Final is very unlikely because [nʌ] is a light syllable)

- At present, this is due to a *negative* weight on ALIGN RIGHT-3.
  - In the Paterian system, this results from the ban on including a final syllable in a foot: [pʌmɛl(ɒnʌ)]
- Step II: refine the model to include a constraint for prestressing suffixes like *-ic*
- Previously, like *acidic*, *atomic*, *synchronic* made it *look like* English is indifferent to the Latin Stress Rule (which wants them to have antepenultimate).
- We include this, arriving at the current state of the model:

<b>-1.96</b>	<b>1.92</b>	<b>1.87</b>
AlignR-3	AlignR-4	AlignR-5

pæ	mɛ	lɒ	nʌ
0.058	0.825	0.116	0.001

- The forms that shift the percentages can be seen in hypothetical *pamalanic* ‘pertaining to Pamelana’:

pæ	mɛ	lɒ	nɪk
0.007	0.097	0.896	0.000

## 21. Upshot

- We cannot easily arrive at an experimental prediction like “participants should favor antepenultimate stress in L L L words”.
- We may need to include other stuff that influences the outcome:
  - *-ic*
  - \*STRESSED HIATUS: *ˈNarnia* vs. rare *Euˈboeia*



- verbs: a two-syllable window when unsuffixed: normal *de'velop* vs. rare *'jettison*
- And then we are probably obliged to blink-test these elements too, to see if they really would override the default.
- ... we must be very careful about this, since we may be throwing away the chance to explain experimental data (below)

## 22. Another example

- Conjoined constraint from last time:
  - “Don’t be a verb, end in VV(C) or VCC, and have non-final stress”
- We had sought to “atomize” the constraint:
  - “Don’t be a verb and have non-final stress”
  - “Don’t end in VV(C) or VCC, and have non-final stress”
- This was working pretty well: both components got appreciable weights and the conjoined constraint got a small weight of 0.52.
- Now, with more constraints in the grammar (and accuracy up), the weight of conjoined is 1.07.
- Perhaps we should have expected this, for the conjoined constraint has a diachronic explanation:
  - *ar'rest* was once *ar'restən* or something like that, with a heavy penult.

## 23. Generality vs. accuracy

- Moore-Cantwell’s primary Weight-to-Stress constraint applies across the board.
- As she shows, splitting it up by position (e.g. antepenult, penult, etc.) improves fit to the lexicon.
- Hence, opting for generality sacrifices accuracy — more on this below.
- The detailed model seems unprincipled (unless we adopt Liberman-Prince ternary feet, banning foot-medial heavies?)
- Accuracy is fun to pursue (hence my own model, ever aiming for another couple percentage points), but the long term goal is matching native speaker intuition and here generality may be crucial (M-C readings)

## SOME FURTHER CONSTRAINTS I ADDED

## 24. Constraints that reference specific morphemes (affixes)

- These show, at some level, that stress is not predictable.
- But if the affix is reasonably common, they have some generality.

## 25. A criterion for affix-specific constraints

- We don’t expect them to introduce totally bizarre phonology; they “fit in” gently, somehow.

- See Liberman and Prince, who link affixes to the subscripted parentheses in their stress rule.

## 26. RESPECT THE STRESSED SUFFIXES

- Source: Liberman and Prince (1977)
- The stressed suffixes of English tend to be of French origin and some are indeed not quite phonologically integrated.
- Here are the suffixes I included:

*-esque*

*-ette, -aire, -ique, -isce, -enne, -esse, -elle, -aise, -esce*

*-eer, -ier, -ese, -eur, -air, -oon*

*-ee, oo*

- Still to go in: *-age* when pronounced [-ɑʒ]
- Weight of RESPECT THE STRESSED SUFFIXES: 4.79
- Exceptions as follows.

### Falsely parsed suffixes (errors)

cop-ier, fl-ier

di-ocese

### Nativized words

coffee, levee, squeegee, toffee, committee, filigree

cuckoo, igloo, voodoo, cockatoo

amateur

espalier

etiquette

### Change in progress

entrée

chimpanzee

divorcee

fiancee

portuguese

jubilee

waterloo

## 27. A source of permanent non-predictability

- *Degree* of nativization of loan words.
- This has been going on for centuries; e.g. difficulty of scanning Chaucer, for whom many now-fully-nativized French words were new.

## 28. Respect stressless prefixes in disyllabic verbs

- I put in:

*de-, re- ad-, ex-, in-/im-  
con-/com-, per-, sub-/sup-, pro-, dis-, mis-*

[ ☞ Are there some I forgot? ]

- These come from Latin.
- They are heavily involved in the noun-verb stress alternation, e.g. *contrast*, *permit*, *conduct*, etc. (hence, grist for the constraint on verb-specific final stress)
- They are involved in Cyclic Effect #6:
  - a deverbal noun generally retains the final stress of its verbal base, as a secondary stress

<sup>1</sup> per <sub>1</sub> mit	vs.	<sup>1</sup> hermit
<sup>1</sup> ad <sub>1</sub> dress	vs.	<sup>1</sup> wondrous
<sup>1</sup> tor <sub>1</sub> ment	vs.	<sup>1</sup> dormant

- Weight of RESPECT STRESSLESS PREFIXES is 2.81
- Exceptions are surprisingly rare:<sup>2</sup>

compass  
conjure(do children think this has final stress?)  
conquer  
convoy  
injure  
perjure  
proffer (etymologically: pro-offer)  
prosper  
revel (does this have final stress?)

## 29. Dealing with extrametricality

- As we saw, some suffixes (and the ends of a few monorphemes) behave accentually as if their final syllable is “not there”.
- Here are the relevant suffixes:

Suffix	Examples	Base (?)
-ative	<sup>1</sup> generative, <sup>1</sup> imi <sub>1</sub> tative, af <sup>1</sup> firmative, pe <sup>1</sup> jorative	<sup>1</sup> generate, <sup>1</sup> imitate, * <sup>1</sup> pejorate
-ancy	<sup>1</sup> hesitancy	<sup>1</sup> hesitant

<sup>2</sup> I ignore cases where orthography misleads us, like *redde*n

-ator	'agitator, 'narrator	'agitate, 'narrate
-able	in'heritable, 'malleable	in'herit, *'malleate
-acy	le'gitimacy	le'gitimate
-ary	'budge,tary, parlia'mentary, pe'cuniary	*pe'cunia
-ory	'circula,tory, 'cursory	'circulate, *cursate
-ism	'despo,tism, ca'tholi,cism, 'mechanism	'despot, 'Catholic, mechan(ist)

- The empirical effects are:
  - a larger stress window ('generative)
  - choice of other than rightmost final main stress ('despo,tism)
- They often seem largely explicable as inheritance of the stress of a shorter stem.
  - I like this idea because it has been fruitful elsewhere (Steriade, Breiss on *-able*)
  - but it doesn't always work for extrametricality; see above
- sometimes the phonological default asserts itself: 'explicate, ex'plicable
- For now, I just assume a “pseudo-base”, constructed by removing the last syllable.
- This improves model performance a lot, since it covers most of the exceptions to ALIGN RIGHT-4 and ALIGN RIGHT-5, enabling them to be boosted in weight.

### 30. This will not solve all the problems with “rightmost nonfinal main”

- I.e. there are a fair number of *monomorphemic* items that behave similarly:
- These, amusingly, sometimes end in a sequence that could a suffix:  
*cassowary, alligator*

Other:

*axolotl, Aristotle, filibuster, alabaster, orthodoxy, apoplexy, difficulty, controversy, epilepsy, testimony, parsimony, patrimony, matrimony, acrimony, alimony, ceremony, ignominy, taxidermy, melancholy, patriarchy, matriarchy, oligarchy, demagogu, pedagogy, necromancy, helicopter, oleander, necromancer, architecture, nomenclature, cataclysm, microcosm, enthusiasm, logarithm, participle, tabernacle*

FTBIN

### 31. The constraint

- \*A foot with just one mora.
- Mostly statable observationally:
  - No stress on a light syllable that is either word final or directly before another stress.
- This is an important element both in Pater's analysis and in the Moore-Cantwell MaxEnt analysis discussed below.

- FTBIN is the modern account of most of what SPE and other rule-based work modeled with “Destressing” rules: *phonetician*, *acidic*.
- Moore-Cantwell, in her MaxEnt model: “FTBIN gets the highest weight of all the constraints, since it is essentially unviolated in the training data.”

## GROVELING IN THE LEXICON FOR FTBIN EXCEPTIONS

### 32. Final position

- *yeah*, *duh*, which are gestures not words (cf. [ˈʔΛʔou], [ˈʔΛʔΛ], [ˈjɛ], [ḡhḡm])
  - And indeed, perhaps they have long vowels
- Dr. Seuss’s novel letter *Nuh* (from *On Beyond Zebra*, used to spell *Nutches*).
- None at all in polysyllables: \*[fəˈdΛ], \*[pæləˈmɛ]

### 33. \*Pretonic 2ary-stressed light

- These are found in the initial syllables of some French/foreign-tinged disyllables, with [æ], as in *ballet* [bæˈleɪ], also

*abbe*, *cachet*, *cafe*, *caffeine*, *cashier* (Level II?), *chateau*, *malaise*, *massif*, *passe*, *plateau*, *raccoon*, *rapport*, *rattan*, *sachet*, *taboo*, *tattoo*, *valet*, *vasectomy*, *paprika*

These may joined by *settee* and *suttee*, with mid vowels; none with high (\*bɪˈleɪ)

- Words where, I think, people can optionally suppress vowel reduction in careful speech, especially when the word is felt to be rare or fancy. I feel these are actually stressless ...

*cessation*, *effete*, *enable*, *ennui*, *essay*, *penultimate*, *premier*, *stenographer*, *bitumen*, *chinook*, *Nippon*

- Again, “gesture-not-word” words: *yippee*,<sup>3</sup> *hello*

### 34. The ban on weak lights and Maximal Onset Syllabification

- The following look like MOS-violating assignments of coda position in s+stop clusters:

*cascade*, *chastise*, *domesticity*, *elasticity*, *escarpment*, *escort*, *espalier*, *estimator*, *existential*, *existential*, *fastidious*, *festivity*, *festoon*, *gestation*, *hysteric*, *manifestation*, *mascara*, *mastectomy*, *mestizo*, *molestation*, *pastel*, *pastiche*, *prestige*, *prestigious*, *protestation*, *risque*, *trustee*

- This makes me wonder if STRESS-TO-WEIGHT causes resyllabification of the /s/ into coda?

<sup>3</sup> Mid 20th century English for “hooray”. Adam Albright points out somewhere that the [p] is not aspirated.

### 35. Monomoraic feet also occur, not commonly, in main stress position

<i>ally</i>	AE1	543
<i>asset</i>		398
<i>rabbi</i>		123
<i>addict</i>		109
<i>fascism</i>		109
<i>alloy</i>		40
<i>adjunct</i>		33
<i>annex</i>		31
<i>graphite</i>		15
<i>cashew</i>		9
<i>atoll</i>		7
<i>cathode</i>		7
<i>anode</i>		1
<i>decade</i>	EH1	1394
<i>essay</i>		405
<i>telex</i>		83
<i>methane</i>		78
<i>semi</i>		21
<i>serrated</i>		19
<i>narrate</i>		13
<i>reggae</i>		8
<i>Semite</i>		5

- A few more under maximal-onset syllabification (I leave out s+stop):  
*tabloid, attribute, address, padre, refuse, jaguar, legume*
- Observe again the preference for lower vowels, though here [ɛ] is better represented.  
There is an interesting repair strategy for such feet: the following syllable can be destressed ...

### 36. How well is Bruce's maxent model performing?

- Probability that best guess is true: This is currently **.881**.
- Average probability assigned to best guess  
➤ This is currently **.805**.

### 37. Log likelihood

- The gold standard, for many, is log likelihood of the data.
- At present: -4485.

### 38. Finding meaning in log likelihood

- Let's divide by total forms:  $-4485/12,664 = -.354$
- This is the average penalty in log likelihood.
- It corresponds to  $e^{-.354} = .702$
- This is the *geometric* mean, as opposed to the arithmetic mean above.
- Diagnosis: it responds to “stinker” forms, where the winner gets a very low probability.

### 39. The 20 worst stinkers (☞ why?)

brutalization	B R UW1 / T AH0 / L AH0 / Z EY2 / SH AH0 N	0.000092
formalization	F AO1 R / M AH0 / L AH0 / Z EY2 / SH AH0 N	0.000092
deterioration	D IH0 / T IH1 / R IY0 / ER0 / EY2 / SH AH0 N	0.000191
alternation	AO1 L / T ER0 / N EY2 / SH AH0 N	0.000586
circulation	S ER1 / K Y AH0 / L EY2 / SH AH0 N	0.000586
convolution	K AA1 N / V AH0 / L UW2 / SH AH0 N	0.000586
decimation	D EH1 / S AH0 / M EY2 / SH AH0 N	0.000586
destitution	D EH1 / S T AH0 / T UW2 / SH AH0 N	0.000586
immolation	IH1 / M AH0 / L EY2 / SH AH0 N	0.000586
syncope	S IH1 NG / K AH0 / P EY2 / SH AH0 N	0.000586
igloo	IH1 / G L UW0	0.000703
inventory	IH2 N / V AH0 N / T AO1 / R IY0	0.001135
excoriation	EH2 K / S K AO1 / R IY0 / EY2 / SH AH0 N	0.001217
incineration	IH0 N / S IH1 / N ER0 / EY2 / SH AH0 N	0.001217
apropos	AE2 / P R AH0 / P OW1	0.001408
clarinet	K L EH2 / R AH0 / N EH1 T	0.001408
entourage	AA2 N / T UH2 / R AA1 ZH	0.001408
esplanade	EH2 / S P L AH0 / N AA1 D	0.001408
harmattan	HH AA2 R / M AH0 / T AE1 N	0.001408
minuet	M IH2 N / Y AH0 / W EH1 T	0.001408

## EXPERIMENTS ON ENGLISH STRESS

### SOME GENERAL REMARKS ON EXPERIMENTAL PHONOLOGY

#### 40. The forms of test

- Wug (Berko 1958)
- Blick (Chomsky and Halle 1965)
- Stride (Kuo 2020)

**41. What would a stress test be?**

- It is blick, since the placement of stress is part of the phonotactic system.
- It is wug, because the placement of stress on a syllable represents a choice (like an affix allomorph in vowel harmony).

**42. Four works of experimenting on English stress I've looked at**

- Larry Nessly (1977) in *USC Occasional Papers in Linguistics*
- Trammell (1978) in *Journal of Psycholinguistic Research*
- Guion et al. (2003)
- Moore-Cantwell (2015, readings)

**SOME REMARKS ON NESSLY****43. The field in 1977**

- He notes that very few phonologists did experiments; this lasted through the 20th century; experiments were the domain of theory-haters!

**44. Orthography**

- He buys into SPE's view that orthography is very close to UR; hence he tests whether orthographic geminates attract stress.

Blick words: ocellus, ocelus

**45. He knew about the sonorant/obstruent generalization, roughly**

caroptove should be penultimate stress, and more or less does  
miftenpet should get antepenultimate stress, and more or less does  
miftempet should be penultimate stress and more or less does

**46. A cute fact**

- The sonorant/obstruent generalization overrides orthographic gemination!
- ocellus vs. ocellote

**47. The foreign-word stress rule**

Stress the penult when the final is CV, else the final.

He had participants in both Ann Arbor and Austin, and the Texas applied the foreign rule more often; he thinks due to great proximity to Spanish.



**48. Things we wouldn't do today**

Ask in person (Clever Hans)  
Not check for near neighbors  
Present orthographically (probably not)  
Very primitive stats (Fisher's exact test only)

**SOME REMARKS ON TRAMMEL****49. This is (sort of) a stride test**

- English has tons of words that look real, but which hardly any experimental subject knows!
- paludism, hegumen, opodeldoc, etc.

**50. Subject agreement**

- Reassuringly, subjects agree with each other 87% of the time on main stress, but probably a lot of models could achieve this.
- They varied much more on secondary stress and segmentals.

**51. Etymological source was varied**

- Germanic, Latin, Greek
- No effects found
- This seems premature to me — we first need a model of how speakers might assign to strata (see Hayes 2016 CLS for a MaxEnt attempt)

**52. Alas, orthography doesn't give vowel length**

- Simple examples like *monip* tell us this.
- Vowel length can be semi-predicted, e.g. by watching closed syllables, geminates.
- But it's a severe drawback in stress experimentation

**GUION ET AL (2003)****53. Innovation #1**

- Play the syllables in isolation, then two plausible concatenations.
- either produced by the subject, or multiple choice
- Yay, a solution to the orthography dilemma

**54. Innovation #2: theory**

- "Phonological models would do well to accommodate multiple and competing sources of information leading to predictions about stress placement."

- I.e. abandon the “one derivation” system that dominates most of the history of English stress.
- ... and they have logistic regress a.k.a. MaxEnt to do this for them

## 55. Neighbor words

- These are big in analogical models.
- There is an effect, but a modest one.

MOORE-CANTWELL (2015)

## 56. This one is totally up to date — do it this way, guys!

- Guionian presentation of stimuli
- Mixed-effects regression to interpret results
- Results checked against an authentic phonological model.
- UG bias is pondered.

## 57. Looking at the results on experiment with a model in mind

- A probabilistic model offers a more refined measure of fit and can accommodate cases where the participants felt (collectively) that more than one outcome was ok.

## 58. M-C’s model

- It occupies a fourth slot in our taxonomy of models

	Single-solution nonstochastic	Stochastic
Ingredients given by a UG	Pater (2000)	Moore-Cantwell 2015
Ad hoc ingredients allowed	SPE, Liberman and Prince	Hayes, this course

- Fairly Paterian in its constraints, MaxEnt trained on the Hayes-groomed CMU corpus.

## 59. A UG test

- There are very few words like *galaxy*, with closed penult and fitting the M-C [i] generalization.
- Hence a theory needs to project an answer for such cases, if speakers agree with one another.

## 60. Testing the Guionian view

- The MC-I constraint and the heavy penult principle turn out to be independent.

### **61. Defending stochastic grammar**

- Look at each subject separately — they vary like they do collectively.

WHY DO PARTICIPANTS UNDERMATCH THE LEXICON IN VARIOUS RESPECTS?

### **62. A standard research issue**

- Deviations from frequency-matching as resulting from UG principles

### **63. MC's proposed bias**

- Simplicity (well supported elsewhere)
- a heavy penult bias confined to penults would fit the data better
- she suggests speakers have position-free WEIGHT TO STRESS

### **64. Other stuff Bruce ponders**

- Which words they used to acquire the system
- Is there a way to model inattention/indifference/stupidity/malevolence? I.e. can we model our results that fall short of frequency matching by modeling the task?

### **65. What else?**

- If we put together an obsessively detailed lexicon-matching model, would its details match with experimental results?