1. Assignments

- Read for next time:
  - I will mail it out; web site still down.
- Homework on bias due in class Monday April 23.

2. Looks good

- Everyone is with the program, and various imaginative things popped up.

3. Integrating more traditional phonology

- The traditional story is define onsets, define codas, concatenate, add a few trans-syllabic constraints.
- This would imply including the initial and final clusters in the spreadsheet, trying general constraints that penalize e.g. both $CC$ and $C#$.
- Then, we have the problem of syllabification; is [b] in [abra] an onset or coda?
  - The problem of *hidden structure*, to be covered later.

4. Repaired violations

- Sometimes a trans-syllabic constraint is actively repaired, so the constraint you posit interact with Faithfulness.

5. The Syllable Contact Law

- Useless. Every single language.

6. A reference source on modeling I couldn’t find earlier

  - Fisheries bit was: Colorado Cooperative Fish and Wildlife Research Unit.
- Coetzee and Pater cite an article by this helpful pair of statisticians.
MORE ON KNOBS

7. What have we got so far?

- Four conjectured “knobs”:
  - style
  - emphasis
  - rate
  - lexical frequency
- A fifth knob made dubious by acquisition issues:
  - identity, depending on exposure to enough learning data in childhood
- Classical methods of measuring the effect of style: Labov’s interviews and observations
- Lockstep: New Yorkers lockstep their application of:
  - ɹ → ∅ in codas
  - /æ/ → [iə] before certain consonants
  - /ɔ/ → [ʊɔ]
  - /θ,ð/ → [t̪θ, d̪ð] or [t̪, d̪]
- We were about to examine some lockstepped Lower East Siders

8. Phonological free variation in the speech of Miriam

- Miriam is 35 years old, graduated Hunter College and St. John’s law school, works as lawyer.

- The phenomenon that an upstate-raised General American speaker (BH) finds eeriest is theta-hardening, where Miriam is at zero in all styles.

---

1 I may have paid this short shrift: while I cannot talk like a New Yorker or British person, their accents are easily intelligible to me, unlike (say) obscure rural Irish dialects. I clearly have learned something for purposes of perception, perhaps even post-critical period. See later on breaking competence into production vs. perception components.
9. Variation in the speech of Doris

- Doris is 39, homemaker, African-American.
- She doesn’t have perfect lockstep
- Labov thinks that for Doris, and others, r-dropping is more sensitive to style than other processes.

10. Why is Doris not lockstepped? A conjecture

- Perhaps she is bidialectal in African-American Vernacular English?
- Conceivably she is switching dialects as well as styles?

11. Is there more available somewhere?

- Sociolinguistics seems to be is shifting its emphasis away from phonology …
- But the older sort of data — careful tracking of application rates of multiple processes across style-controlled elicitation — seems our best hope for studying the style knob.

12. Another way to check lockstep: two processes in the same word

- *Tantalus* /ˈtæntələs/. Rare, (see below), which encourages non-Tapping.
- Eligible processes:
  - NT Tapping: \{nt\} → ŭ / V [V ːstress]
  - /æ/ Tensing: æ → ɪə / ___ {m,n}
- Candidate pronunciations:
  - [ˈtɛntələs]
  - [ˈtɪŋtələs]
  - [ˈtæŋtələs]
• ['tiðələs]

- Socrates: what is going on here?

13. Here is another example

- mountain /ˈmaʊntən/.
- Eligible processes:
  - NT Tapping: \{nt\} → ř / V  
    Syllabic Nasal Formation: ən → n / {t,d}
- Candidate pronunciations:
  - ['maʊntən]
  - *[ˈmaʊɾ̃n̩] impossible, Tapping requires a right-side vowel
  - ?[ˈmaʊɾ̃ən] I don’t like this, but others can say it. ²
  - ['maʊntŋ]

14. Term paper topic?

- Formalize the Coetzee/Pater knob below and make correct predictions about such cases, sometimes known as “register conflict”.

15. Free variation in society is structured as well

- Fig. 4.2 from William Labov (1972) Sociolinguistic Patterns

---

² My evidence is a recording by the folk singer Pete Seeger, listened to by me on disk in childhood and cassette in parenthood: “She’ll be comin’ round the ['maʊɾ̃ən] when she comes”.
- from 81 native speakers of New York City English
- Vertical axis: what percentage of underlying /æ/ are retained in the output?
- An independent investigation sorted the speakers into their social classes.
- The “leaping up” of the lower-middle-class speakers in the formal styles is found in other studies, and is claimed to reflect a social insecurity independently diagnosed by other tests.³
- Perhaps these speakers have similar grammars but habitually set their style knobs differently?
  ➢ Again, the “who listens to whom during acquisition” issue arises; lawyer Miriam may have never met Bennie, Labov’s truck driver.

16. The research challenge

- Do knobs exist as entities, so that multiple processes really do vary in lockstep?
- If so, how can we implement knobs in a formal phonological grammar?

³ For example: series of questions: “how do you say this word? … how should this word be said?”, total cases of difference.
Can Harmonic Grammar/maxent help?
- In empirical work, are there rigorous ways for us to track how knobs are set?

A HALLMARK OF KNOBS IN MAXENT: SIGMOID CURVES

17. Goal
- Understand the qualitative predictions a theory makes, so we know what to look for.

18. Scenario
- We set in conflict a Markedness constraint and a Faithfulness constraint, for R Dropping:
  - *CODA r
  - MAX(r)
- We follow Coetzee and Kawahara, below, in assigning our knob to the Faithfulness constraints.
- Under this scheme, MAX(r) gets an augment or decrement, based on setting of some knob.
  - Call it K.
- Harmony of /kar/ → [ka]: weight of MAX(r) + K
- Harmony of /kar/ → [kar]: weight of *CODA(r)
- Let’s do a little spreadsheet, checking every value of K from −5 to +5.

- This is the beautiful logistic curve, a common output pattern of maxent.
- It is centered at K = 0, with the varying slope that reflects the barriers to certainty near 1 and 0.
- The math is presented in full detail in the Supplementary Materials to:
Linguistics 219, Class 6 (4/18/18)  Knobs II

- I suspect that Noisy Harmonic Grammar (below) would behave the same but I’m not sure.

19. A consequences of maxent for knob theory

- If the knob is a simple number (harmony adjustment), it should have
  - small consequences for processes that are at extremes: near-impossible, near-obligatory
  - large consequences for processes that apply with close to 50/50 probability.
- We might try to read confirmation/disconfirmation off of Labov’s diagrams for Miriam, Doris, etc. [term paper topic?]
- Or compare near-obligatory regular Tapping with the less-likely-to-apply NT Tapping, seen in *center* [ˈsɛr̩ə].

EFFECTS OF FREQUENCY IN PHONOLOGY:
NOT ALL OF THEM NECESSARILY GRAMMAR

20. Acquisition effects

- The rare is hard to memorize.
- Hence irregular forms tend to get regularized when frequency goes down.
- See readings p. 81: Bybee showed that old irregular pasts, like *chide* ~ *chid*, got regularized in this way.4

21. Nativization effects

- A foreign word becomes more common in usage.
- It starts to feel ever more strange to give it its faithful foreign rendition.
- Thus, regularization.
  - Partly removing marked, foreign configurations
  - Partly just making more faithful to the orthography
- In my lifetime, I suspect, these have been accommodated:
  - *croissant* [kɹwɑsə̃] → [kɹˈsɑnt]
  - *cappuccino* [kæpʊˈtsɨnoʊ] → [kɛpəˈtsɨnʊ]
  - *gazpacho* [.gasˈpatʃoʊ] → [.gasˈpatʃoʊ]
- I believe that the second example in the readings is of this kind; Japanese speakers getting used to foreign words with [bb], [dd], etc.
- Is there a “perceived foreignness” knob?

---

22. **Online production effects of frequency (focus here)**

- Psycholinguistics has shown that the listener knows the frequencies of words, and weights their probabilities in perception.
  - This is probably one of the best-established results in the field.
- Greater Faithfulness in speaking gives your hearer a better chance on rarer words — this accords with my commonsense experience.

23. **Example: let us experience our Sprachgefühl for Tapping**

- This is from my little utility; words from CMU Dictionary; frequencies from CELEX:\(^5\)
- Obviously-affixed words excluded.

<table>
<thead>
<tr>
<th>Word</th>
<th>Frequency</th>
<th>Word</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>little</td>
<td>21886</td>
<td>gratis</td>
<td>5</td>
</tr>
<tr>
<td>water</td>
<td>8418</td>
<td>natter</td>
<td>5</td>
</tr>
<tr>
<td>matter</td>
<td>6622</td>
<td>bonito</td>
<td>4</td>
</tr>
<tr>
<td>society</td>
<td>6066</td>
<td>carotid</td>
<td>4</td>
</tr>
<tr>
<td>city</td>
<td>4867</td>
<td>catacomb</td>
<td>4</td>
</tr>
<tr>
<td>letter</td>
<td>3706</td>
<td>Catalonia</td>
<td>4</td>
</tr>
<tr>
<td>total</td>
<td>2520</td>
<td>catamaran</td>
<td>4</td>
</tr>
<tr>
<td>committee</td>
<td>2503</td>
<td>clematis</td>
<td>4</td>
</tr>
<tr>
<td>quality</td>
<td>2369</td>
<td>cuticle</td>
<td>4</td>
</tr>
<tr>
<td>daughter</td>
<td>2305</td>
<td>betel</td>
<td>3</td>
</tr>
<tr>
<td>hospital</td>
<td>2300</td>
<td>philately</td>
<td>2</td>
</tr>
<tr>
<td>attitude</td>
<td>2246</td>
<td>poinsettia</td>
<td>2</td>
</tr>
<tr>
<td>pattern</td>
<td>2213</td>
<td>vibrato</td>
<td>2</td>
</tr>
<tr>
<td>bottle</td>
<td>2181</td>
<td>yetì</td>
<td>2</td>
</tr>
<tr>
<td>pretty</td>
<td>2125</td>
<td>lotto</td>
<td>1</td>
</tr>
</tbody>
</table>

**COETZEE AND KAWAHARA**

24. **Coetzee and Kawahara’s hypothesis: the Frequency knob is implemented by Faithfulness variability**

\[
H(cand) = \sum_{i=1}^{n} (w_i + nz_i) M_i(cand) + \sum_{j=1}^{m} (w_j + nz_j + sf) F_j(cand)
\]

- This is the Harmony formula, which they then plug into the Noisy Harmonic Grammar framework, similar in results to maxent.
- The formula sums up Markedness and Faithfulness.
- Since it’s NHG, we have noise \((nz)\), added into every constraint weight.
- The scaling factor \(sf\) (my K) is the same for all Faithfulness constraints (lockstep), and its value comes from another equation relating it to frequency.

\(^5\) Actually, *Britain* and *British* are way up there; I removed them as unlikely to be super-frequent for non-Brits.
25. The first stage of the math: computing the scaling factor from frequency

- Find the value at which this function (called the beta distribution) achieves its maximum:

\[
f(x, \alpha, \beta, \rho) = \frac{\rho x^{\alpha-1} (1 - x)^{\beta-1}}{\int_0^1 x^{\alpha-1} (1 - x) \, dx}
\]

where
- \( \rho = \) general amount of oomph the factor provides; C+K obtained a value by fitting to their data
- \( \alpha = \) log of “reference frequency”; the median frequency of the corpus sorted by tokens
- \( \beta = \) log of frequency of the word you are dealing with

26. Understanding this formula

- I am at a loss for why they chose it.
- The rough qualitative pattern is certainly appropriate:
  - when a word is frequent, the scaling factor is negative (Faithfulness weights go down)
  - when a word is infrequent, scaling factor is positive (Faithfulness weights go up).

27. Graph showing this
28. **Interpreting the scaling factor**

- Noisy Harmonic Grammar: Gaussian distribution for the constraint weight moves sideways by whatever the factor is.
- This changes probabilities of output candidates, in a way that could be solved analytically, I guess, but I will not try here.
- *If* we did it in maxent, our earlier theorem applies: log odds of Unfaithful candidate to Faithful candidate shift by the scaling factor.
- But since this is a two-stage process, we do not have some straightforward equation relating lexical frequency to rule application frequency.

29. **First empirical study: /t/ deletion**

- Buckeye Corpus
- Markedness and Faithfulness constraints are straightforward:
  - Constraints used:

  *CT]*Word
  Assign one violation mark for every word that ends in the sequence [-Ct] or [-Cd].

  MAX
  Assign one violation mark for every input segment lacking an output correspondent.

  MAX-PRE-V
  Assign one violation mark for each segment that appears in pre-vocalic context in the input, and that does not have a correspondent in the output.

  MAX-PRE-PAUSE
  Assign one violation mark for each segment that appears in pre-pausal context in the input, and that does not have a correspondent in the output.

30. **Knob**

- All three Faithfulness constraints get the same knob-based boost/decrement.
- Note that since there is Faithfulness overlap there will be a *double* boost/decrement.

31. **Incorporating frequency improves model performance**

- They use the “AIC” (Akaike Information Criterion) and it unambiguously shows an improvement, properly taking into account the increase of 1 in parameter count.

32. **Commentary I: choice of data**

- The word *and* is extremely frequent and is an outlier in the distribution.
• Indeed, the raw distributions are narwhal-shaped.6

• So to what degree is the model performing as well as it is because of and?

33. Commentary II: testing the model

• I would see the “acid test” to be taking on some “lockstep” data, as above. Perhaps one phenomenon at a time is too easy?

GENERAL PREDICTIONS OF THE MODEL

34. Prediction (p. 78): no frequency reversals in different Markedness contexts

• If deletion word-finally is more common than deletion pre-consonantally for some lexical frequencies, it must be so for all lexical frequencies.
• Same, I suppose, for other knobs.

35. Prediction (p. 80): Markedness conflicts do not respect word frequency differences

• Stress patterns are often expressed with a set of conflicting Markedness constraints.
• There is attested free variation in stress patterns; e.g. penult/preantepenult in LLLL words of Egyptian Radio Arabic.
• So these should not be sensitive to lexical frequency.

36. Restating the prediction

• The standard treatment of allophones in OT is ranking of Markedness constraints only.
• The very fact that they are allophones means that the lexicon has no influence.
• Hence Faithfulness excluded.
• This excludes all instances of allophony from word-frequency effects — a bold move!

6 Thanks, Beth and Connor, for this outstandingly useful term! I suggest defining it as “blob, plus outlier creating a good correlation”
• This ought to be checkable against the research literature in phonetics.

37. Can we hand-check the role of frequency in an allophonic process?

• Let’s try my (and perhaps your?) English, with diphthongization of /æ/ to [iə] before [m] and [n].
• Frequencies from CELEX:

<table>
<thead>
<tr>
<th>Phrase</th>
<th>Count</th>
<th>Word</th>
<th>CELEX Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>andiron</td>
<td>1</td>
<td>and</td>
<td>514946</td>
</tr>
<tr>
<td>annular</td>
<td>1</td>
<td>can</td>
<td>71194</td>
</tr>
<tr>
<td>anode</td>
<td>1</td>
<td>man</td>
<td>29731</td>
</tr>
<tr>
<td>banditry</td>
<td>1</td>
<td>hand</td>
<td>14241</td>
</tr>
<tr>
<td>pantaloon</td>
<td>1</td>
<td>stand</td>
<td>8954</td>
</tr>
<tr>
<td>aggrandize</td>
<td>1</td>
<td>answer</td>
<td>5435</td>
</tr>
<tr>
<td>manioc</td>
<td>2</td>
<td>plan</td>
<td>5429</td>
</tr>
<tr>
<td>Tantalus</td>
<td>2</td>
<td>land</td>
<td>5152</td>
</tr>
<tr>
<td>galvanic</td>
<td>2</td>
<td>animal</td>
<td>4658</td>
</tr>
<tr>
<td>polyandrous</td>
<td>2</td>
<td>demand</td>
<td>3584</td>
</tr>
<tr>
<td>aniline</td>
<td>2</td>
<td>chance</td>
<td>3221</td>
</tr>
<tr>
<td></td>
<td></td>
<td>plant</td>
<td>2707</td>
</tr>
</tbody>
</table>

• I think I diphthongize the common words more.
  ➢ It sound pretentious to me to say the common words as [æ], perfectly natural to say the rare words as [æ].
• But this actually be the style knob: the rare words are suited to learned and vernacular styles in any event.
• This might at least show that the style knob must manipulate Markedness as well as Faithfulness — allophonic processes are style-markers par excellence.

38. What about fortition?

• Do we not take advantage of “clear speech” allophones when we convey a rare or novel words?
• “Hello, my name is Fred [tʰælfəs]” (Talfus) with strong aspiration of /t/, really low [æ]?
• Like Faithfulness, this too helps the listener by providing salient, dispersed allophones.

39. Prediction: cases violating multiple overlapping Faithfulness constraints should be more frequency-sensitive

• Scenario: optional vowel reduction of unstressed [i, e] is less common for non-high vowels, word-final vowels
  ➢ /ˈpalipu/ → [ˈpaləpu] 80%
  ➢ /ˈpalepu/ → [ˈpaləpu] 60%
  ➢ /ˈpapuli/ → [ˈpapulə] 40%
  ➢ /ˈpapule/ → [ˈpapulə] 30%
• Faithfulness constraints:
  ➢ IDENT(vowel quality) / in non high vowels
  ➢ IDENT(vowel quality) / ___ ]

• Prediction:
  ➢ /papule/ → [papulə] should be more as sensitive to lexical frequency
  ➢ (or to other knobs should one wish to extend the theory)
  ➢ Why? It gets two copies of the Faithfulness knob value entered into its harmony computation.