Class 4, 4/8/15: Model Evaluation; More on Iambic Pentameter

1. Course bureaucracy

- Continu reading Hayes/Wilson/Shisko
- Think about term paper projects — what data would you like to analyze? Make appt.

MODEL EVALUATION

2. Readings

- This lecture summarizes Hayes, Wilson and Shisko (2012: 712-713)

3. Skill at hand

- Use statistical testing to decide what constraints belong in a maxent grammar

4. Sample data

- The Hausa *mutadaarik* meter presented as an exercise by Russ last time.

5. A possible coding of the data

- - / - - / - - / - -	6
- - / - - / - - / v v -	0
- - / - - / v v - / - -	0
- - / - - / v v - / v v -	2
- - / v v - / - - / - -	6
- - / v v - / - - / v v -	1
- - / v v - / v v - / - -	7
- - / v v - / v v - / v v -	0
v v - / - - / - - / - -	1
v v - / - - / - - / v v -	2
v v - / - - / v v - / - -	3
v v - / - - / v v - / v v -	0
v v - / v v - / - - / - -	4
v v - / v v - / - - / v v -	0
v v - / v v - / v v - / - -	3
v v - / v v - / v v - / v v -	1

6. Simplications (shameless, motivated by pedagogy)

- One line beginning v -:

  ∅ Shigaa wuta too sai zaaluncii

  - let's just ignore it for now; it may be meaningful though since other Hausa poetry has lines beginning with "gaps"

- Let us ignore two **false quantities**:

  007b: Yaa kwant**aa** gadoo nai yai barcii should be light
  015a: Da rashin yar**da** da faɗar Allah should be heavy

7. Some constraints assumed to be infinite-weight

- LINES HAVE FOUR FEET
- FEET ARE TETRAPOSITIONAL
- ONE MORA PER POSITION
- LAST TWO POSITIONS OF FOOT MUST CORRESPOND TO A HEAVY

With this, we can have a GEN of just 16 members (4 binary choices)

8. The distribution of the two foot types within the line in the sample

	Foot 1	*Foot 2*	*Foot 3*	*Foot 4*	*Total*
v v -	13	22	16	5	56
- -	22	14	20	30	86

9. My feelings when I eyeballed the data

- The two foot types compete rather equally.
- / - - / is dominant in the last foot.
- Nothing else leapt to mind.

10. Defn. *Quantitative clausula*

- A metrically-invariant or near-invariant sequence at or near the end of a line

  - I learned this term from Roman Jakobson's work (analysis of the Serbo-Croatian epic pentameter)
  - Jakobson, R. 1933. Uber den Versbau der serbokroatischen Volksepen. *Archives neerlandaises de phonetique experimentale* 7-9: 44-53. Reprinted in J akobson (1966) *Selected writings IV: Slavic epic studies*. The Hague: Mouton.

11. A constraint I feel pretty good about (empirically at least)

FINAL CLAUSULA: The line must end in - -.

12. Constraints I'm at least willing to toy with

MEDIAL CLAUSULA: the first half-line must end in v v -.
HEAVY PREFERENCE: weakly prefer - to vv.

13. Setting up the spreadsheet

- Candidates — 16 of them
- Constraints (3) and violations
- Apparatus to calculate Harmony, eHarmony, Probability, and Plog, as before

14. Eight grammars

- Either include, or leave out, each of the three constraints.
- Yielding 8 plogs.

15. The eight grammars sorted by constraints they contain

- - Clausula	HeavyPref	Medial Clausula	log likelihood
yes	yes	yes	-89.18
yes	yes	no	-91.01
yes	no	yes	-90.18
yes	no	no	-91.08
no	yes	no	-97.07
no	yes	yes	-92.80
no	no	yes	-98.92
no	no	no	-99.81

16. The eight grammars sorted by plog

- - Clausula	HeavyPref	Medial Clausula	log likelihood
yes	yes	yes	-89.18
yes	no	yes	-90.18
yes	yes	no	-91.01
yes	no	no	-91.08
no	yes	yes	-92.80
no	yes	no	-97.07
no	no	yes	-98.92
no	no	no	-99.81

- Obviously, the richest grammar has the highest plog.

- Yet it is also the most **complex** grammar — does use of all three constraints make it worth it?

17. The Likelihood Ratio Test

- Use it for comparing *nested* grammars — one grammar has a subset of the constraints of the other.
- Method:
  - Find the difference in plog.
  - Multiply this by two.
  - Look up the result in **the chi distribution** to obtain a p value.
  - Excel: =CHIDIST(2plogdiff, degfreed)
    — 2plogdiff = twice the difference in plog
    — degfreed = **degrees of freedom**, meaning difference in constraint population
  - Reckon the p-statistic as you will: probability that the improvement in grammar performance from adding the extra constraints is due to chance

18. Searching big constraint sets: Top down and bottom up

- Top down: start with the biggest grammar, progressively trim back constraints with the crummiest p-value. Stop when every constraint resists trimback at the significance level you want.
- Bottom up: start with the *null* grammar, progressively add in the constraint that tests with the best p-value.

- Applied to the 87-constraint system used by Hayes/Wilson/Shisko, top-down and bottom up yielded similar but not identical grammars.

19. Spreadsheet: implementing top-down and bottom-up for the mutadaarik

- Upshot: my "gut feelings" were right — only the final clausula constraint seems to be worthwhile.

LESS-OBVIOUS STUFF ABOUT PENTAMETER

20. Working more rigorously

- We linguists ought to be able to find subtler things, drawing our theoretical understanding of phonological structure and our practice in scrutinizing data with great care.
- This actually seems to have happened, mostly with the work of Paul Kiparsky in the 1970's.
  - Stress, syntax and meter (1975), *Language*
  - The rhythmic structure of English verse (1977), *Linguistic Inquiry*

21. "Lexical" stress

- A stress is **lexical** is if is a stressed syllable of a polysyllabic word.
- Lexical stresses are regulated more tightly than other stresses (Kiparsky 1975)

This is clearly a complex line, but not all that unusual in Shakespeare:

```
Pluck the keen teeth from the fierce tiger's jaws,
  w   s    w   s    w   s    w  s w    s
```

This is a type of line Shakespeare virtually never writes:

```
*Pluck immense teeth from enraged tiger's jaws,
   w   s   w     s    w  s  w   s w    s
```

The difference is verified in Hayes/Wilson/Shisko's maxent analysis.

- English does allow lexical mismatches in inversion — always after a phonological break.

Canker'd with peace, to part your canker'd hate. *Romeo and Juliet*

22. Variations on the lexical-inversion theme

- German and Russian verse permit inversion, but not initial *lexical* inversion.

```
                                            x
              x         x                   x
 .   x .      x         x    .   .   x  .   x
Und neuen Glanz ][ schöpf ich aus deinem Schatten
 .   x .      x         .    x   .   x  .   x

                                            x
              x         x                   x
 .   x .      x         x  .   .   x  .     x
*Und neuen Glanz ][ schöpfen aus deinem Schatten
 .   x .      x         . x    .   x  .     x
```

'and new luster I/to create from thy shadow'

- References:
  - — Alan Prince 1989, "Metrical Forms," in Kiparsky and Youmans (1989), *Rhythm and Meter*
  - — Bjorklund, Beth (1978) *A study in comparative prosody : English and German jambic pentameter*

- Milton *every once in a while* mismatches a lexical stress not after a break:[1]

[1] The end of a long discourse; Adam addresses himself, realizing he has committed not just sin but *original* sin, dooming all of humanity.

Thus, what thou desirest
And what thou fear'st, alike destroys all hope
Of refuge, and concludes thee miserable
Beyond all past example and <u>future</u>.

but not at the end of a word, as in "immense teeth" (an endings-strict effect? more later)