Linguistics 219 | Spring 2020
Phonological Theory III | B. Hayes

Class 16, 5/21/2020: Generative Phonetics II, MaxEnt

1. Assignments

- Reading:
  - Michael Lefkowitz (2015) Learning phonetic harmonic grammars for duration. Handout for colloquium at UCLA.
  - On course web site.
  - I have also posted the whole 2017 dissertation if you are curious for more.
- Homework #4 (Indonesian) due next class (Tuesday)
- Feel free to discuss progress/problems with your term project.
  - Office hours W, F at 2.

2. Looking ahead: last three topics

- Juliet Stanton's recent work on the role of contrast in phonology.
- A bit on Gradient Symbolic Representations
- A bit on the Delaware School of Phonology

3. Summarizing last time

- Generative phonetics as a research program.
- Justifying study of acoustic data as well as articulatory
- Arithmetic regularities are present in phonetic data.
- Methods to find them, e.g. varying a knob and plotting paired data
- Targets and interpolation
- MaxEnt phonetics, predicting frequency distributions
- Parabolic penalty functions in constraints
- Lefkowitz's Theorem: grammars made of parabolas predict Gaussian distributions of outputs
- Braver: using "Flemmingian Parabolic Harmonic Grammar" to model near-neutralization
  - The always-compromise prediction

4. This class

- We move on to MaxEnt, modeling output probability distributions.
- First for Braver's Japanese example, then on to Lefkowitz's work on English duration.
- I will also go through a classic earlier paper, Flemming (2001).

MODELING BRAVER'S WORK IN MAXENT

5. The signature case: Japanese vowel length

- The empirical work was in earlier paper by Braver and his adviser Kawahara.
- Example:
  - ➢ /CV/ → [CVˑ], not quite as long as underlying /CVː/.
  - ➢ /fu/ 'gluten' is [fuˑ] alone, [fu ga] with suffix.
  - ➢ /fuː/ 'seal' is [fuː] alone.
- Example:
  - ➢ /fu/ alone wants to be bimoraic to satisfy a word-minimum.
  - ➢ It wants to be shortish to resemble the base form (seen before a suffix, as in *fu ga*)
  - ➢ There is further evidence — from pitch accent — that the suffixed form is the base in Japanese.

6. Another example, showing experimental setup

Sample stimulus set (from Braver & Kawahara 2016)

condition	*orthography*	
a. short, with particle	木もなくしたよ。	ki mo nakushita yo tree also lost DISC
b. short, no particle	木なくしたよ。	ki nakushita yo tree lost DISC
c. long	キーなくしたよ。	kii nakushita yo key lost DISC

7. The theory in outline

- Derived forms are tied to their bases by weak, weighted constraints that penalize differences in phonetic parameters.
- This is done with Harmonic Grammar, with Flemmingian squared derivations from constraints TARGETDUR=x, OOFAITH(Dur).

8. Trying this out ourselves

- The simplified phonetic data from Braver's *Phonology* article:

condition	mean	SD	rounded
unlengthened short (with particle)	54.99	21.89	50
lengthened short (without particle)	124.98	34.91	125
underlyingly long (without particle)	157.45	39.21	150

- Let us try this with MaxEnt, to get a distribution.
- How good is the fit?
- If we manipulate the weights by hand, what changes emerge?
- What would have happened if we used not Flemmingian parabolas, but a linear (absolute-value) penalty for deviation from target?

FLEMMING'S EARLY RESULT ON LOCI

9. Source

- Flemming, Edward (2001) Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* 18:7-44.

10. The scheme

- You have a consonant, like /t/, and a vowel, like /a/.
  - /t/ wants F2 to be 1800
  - /a/ wants F2 to be 1100.
- The system wants to avoid a severe, mousetrap-like transition at release.
- It turns out that both the consonant and the vowel contribute to the compromise.

11. The existing findings of experimental phonetics

- The outcome is always a compromise.
- The target for the consonant — unobservable, inferred — is called its *locus*.
- This was worked out, with equations, by phoneticians Bjorn Lindblom, Harvey Sussman and others.
  - Lindblom, Björn (1963). Spectrographic study of vowel reduction. JASA 35:1773-1781.
  - Sussman, H. M., K. A. Hoemeke & F. S. Ahmed (1993). A cross -linguistic investigation of locus equations as a phonetic descriptor for place of articulation. JASA 94.1256-1268.
  - Sussman, H. M., H. A. McCaffrey & S. A. Matthews (1991). An investigation of locus equations as a source of relational invariance for stop place categorization. JASA 90.1309-1325.
- The standard "locus equation" for consonants as expressed by Flemming:

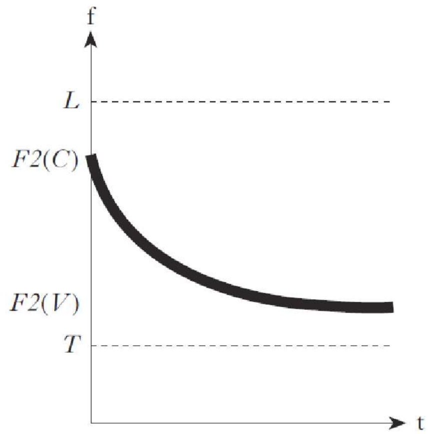
$$F2(C) = k_1(F2(V) - L) + L$$

  - I.e., the F2 of the consonant at release is basically its locus L, but deviating from L by an amount based on how far away the target for the vowel (F2(V)) is.
  - Note that in Flemming's set-up, the vowel itself compromises a bit, deviating from its target.
- Adding in the equation for the vowel:

$$F2(V) = k_2(F2(C) - T) + T$$

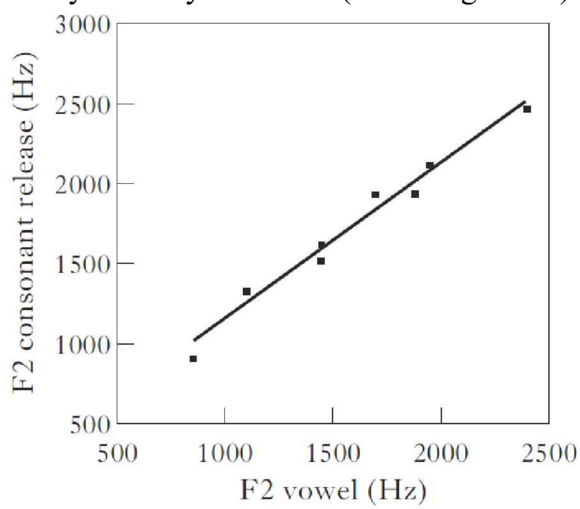
  - which is just the same equation going in the opposite direction.

- Flemming puts the two patterns together in a schematic graph:



12. Locus theory works

- [g] followed by a variety of vowels (Flemming's data):



- ☞ According to Flemming's model (equations above), what is the slope of the line?

13. Flemming's Harmonic Grammar analysis

- Constraints (note squared deviation penalty, per above)

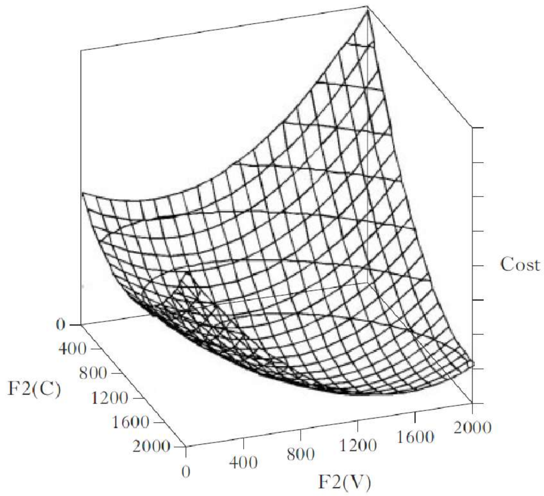
	Constraint	*Cost of violation*
IDENT(C)	$F2(C) = L$	$w_c(F2(C) - L)^2$
IDENT(V)	$F2(V) = T$	$w_v(F2(V) - T)^2$
MINIMISEEFFORT	$F2(C) = F2(V)$	$w_e(F2(C) - F2(V))^2$

- Calculating harmony — this is the standard formula, SUMPRODUCT() with weights.

$$cost = w_c(F2(C) - L)^2 + w_v(F2(V) - T)^2 + w_e(F2(C) - F2(V))^2$$

- Finding the best candidate.

- It is the F2(V) and F2(C) coordinates of the *minimum* of a parabolic harmony-bowl:



- You don't need the Solver for this! Flemming dredged up his high school calculus; a minimum is where the slope along every axis is zero.
- The answer is:

$$F2(C) = u_c(L - T) + L \quad \text{where} \quad u_c = \frac{w_e w_v}{w_e w_c + w_v w_c + w_e w_v}$$

$$F2(V) = u_v(L - T) + T \quad \text{where} \quad u_v = \frac{w_e w_c}{w_e w_c + w_v w_c + w_e w_v}$$

- These equations are indeed the Lindblom/Sussmann locus equations, with coefficients derived from the constraint weights. QED!
- And using the Lefkowitzian math, we could also compute the standard deviations of the Gaussian curves defining the range of variation (are they realistic)?
- ☞ Inspect Flemming's formulae and verify that they behave intuitively under variation of the three constraint weights. See spreadsheet.F

LEFKOWITZ I: EMPIRICAL PATTERNS

14. Source

- Lefkowitz, Michael (2017) Maxent Harmonic Grammars and Phonetic Duration, Ph.D. dissertation, UCLA.
- on course web site

15. The empirical domain: vowel duration in English

- Well studied, so many of the effects he looked at were already documented in the phonetic literature.
- Multiple intersecting factors influence vowel duration, so modeling becomes nontrivial.
- Easy to get data here at UCLA.

16. Some known patterns of vowel duration

- See Lefkowitz's (2017) literature survey for details.
- We can do a quick overview by looking at Lefkowitz's experimental design.
- Varying the segmental context:

	__Ø	__d	__t	__ts			__Ø	__d	__t	__ts
[i]	be	bead	beat	beats			me	mead	meat	meats
[ɪ]	---	bid	bit	bits			---	mid	mitt	mitts
[e]	bay	bade	bait	baits			may	made	mate	mates
[ɛ]	---	bed	bet	bets			---	---	met	---
[æ]	---	bad	bat	bats			---	mad	mat	mats

- Thus:
  - ➢ shorter in closed syllables
  - ➢ perhaps shorter in doubly closed syllables (origin of *keep* ~ *kept*)
  - ➢ shorter before voiced consonants
  - ➢ a little shorter after the nasal (minor effect)
  - ➢ shorter when lax ([ɪ], [ɛ]). We don't know if [æ] is lax.
  - ➢ shorter when low

17. Note: is [æ] lax?

- Traditionally, it is so. Yet
- It can occur in pretonic open syllables: *ballet* [bælˈleɪ], *sashay*, *raccoon*, *Camay*, *Panisse*, etc.
  - ➢ Compare e.g. *[bɛˈleɪ]
- It very marginally can occur finally, e.g. in *baa*, *nah*
  - ➢ Compare *[bɛ], *[bɪ]
- Going the other way: it is fully illegal before a vowel: *[ˈbæoʊ]

18. More known patterns of vowel duration: the prosodic factors

	Unaccented	Accented
Final	Q: Did Bob correctly spell bed? A. No, SUSAN correctly spelled bed.	Q: Did Susan correctly spell fish? A: No, Susan correctly spelled BED.
Medial	Q: Did Bob spell bed correctly? A: No, SUSAN spelled bed correctly.	Q: Did Susan spell fish correctly? A: No, Susan spelled BED correctly.

Table 16: The four prosodic frames, using "bed" as the target word, and "Susan" as the proper name.

- ☞ What is this aspect of the experiment testing?

19. The panoply of good experimental practices

(Be glad, everyone, that you're at a serious phonetics department …)

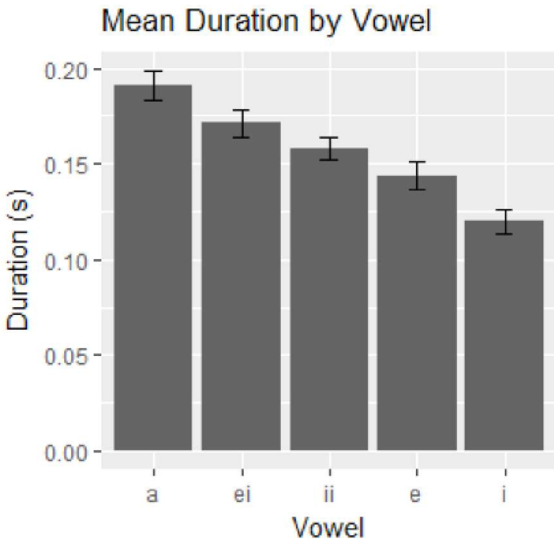
- Distractor sentences
- Sound booth and other elements of good recording quality
- Forced-aligner segmentation for objectivity
- Dual-judge rating for exclusion criteria (wrong phrasing, wrong accentuation, dysfluency)

20. A lesson learned: Many people cannot read aloud fluently

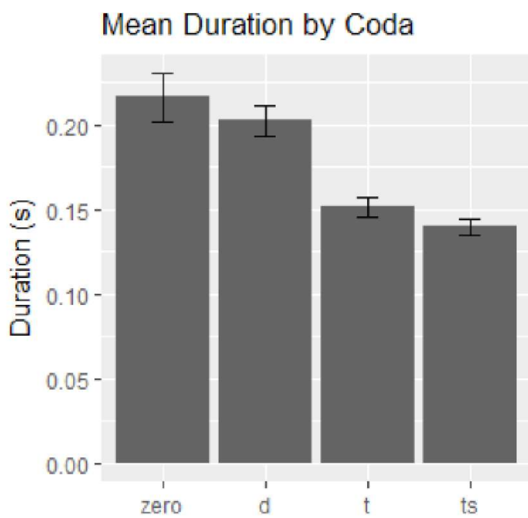
"Perhaps unsurprisingly, most of the unusable items were came from these same subjects: for many, reading aloud in a natural way proved quite difficult, especially when it came to prosodic focus. Many subjects chose to accent all words across the board, put phrase-level juncture between every pair of adjacent words, or both, treating the stimuli more like lists of words rather than sentences. Some failed to produce intonation patterns that were even remotely English-like, and instead seemed to have embarked on random walks through the pitch / duration / intensity space. Others had relatively more natural intonation, but had trouble reading aloud without frequent false starts or segmental dysfluencies."

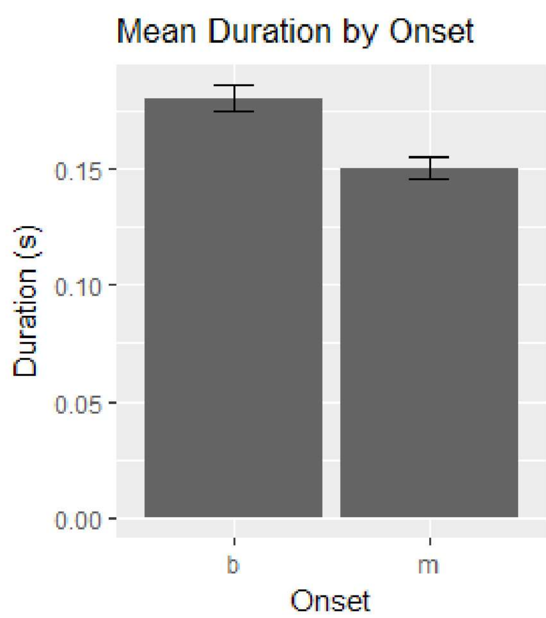
- So subjects who produced more than a criterial number of individual errors had their data thrown out entirely.

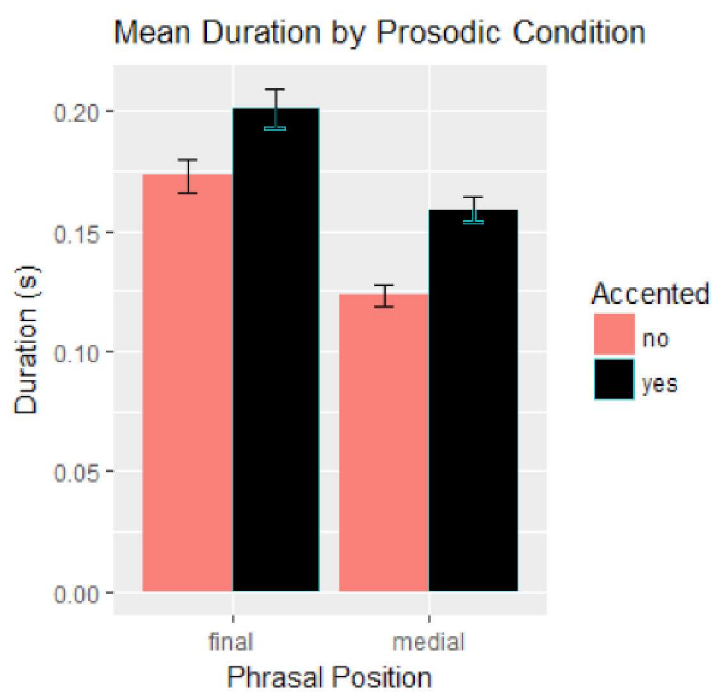
21. Everything in the experiment confirmed conventional wisdom



æ eɪ i ɛ ɪ







22. The point of replicating

- The real importance is in seeing how all these factors **interact**.

LEFKOWITZ II: THE SEARCH FOR TESTABLE QUALITATIVE GENERALIZATIONS

23. No use: Lefkowitz's Theorem

- Recall from before:
  - Parabolic MaxEnt Harmonic Grammars generate Gaussian distributions.
- These distributions are *not* characteristic of phonetic duration patterns, which tend to be asymmetrical about their peaks.

24. Lefkowitz's response

- Complicate the Flemmingian model, going from parabolas to **hemiparabolas**.
- The two constraint types are called:
  - STRETCH: Don't be too long.
  - SQUEEZE: Don't be too short.[1]

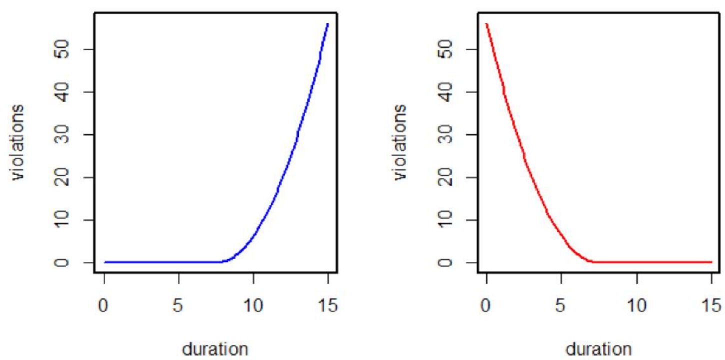


Figure 21: Hemiparabolic violation functions for STRETCH (left) and SQUEEZE (right), with weights set to 1 and targets set to 7.5.

- Hayes and Schuh (2019), working on similar modeling with a far less ambitious dataset (duration in the syllables of sung Hausa poetry) likewise found they needed to move beyond the pure Flemmingian model and use STRETCH and SQUEEZE.
  - (2019) Bruce Hayes and Russell Schuh. Metrical structure and sung rhythm of the Hausa rajaz. *Language* 95: e253-e299 (*Phonological Analysis*).

25. Looking ahead: Lefkowitz's best model for his data

- By "best", I mean a combination of accuracy and few parameters, measured with the Akaike Information Criterion (AIC).
- This model has
  - just one general STRETCH constraint
  - a general SQUEEZE constraint
  - SQUEEZE constraints for every environment studied
- The targets are expressed in deciseconds.

[1] Do I have this backwards? Should the names be *STRETCH and *SQUEEZE?

constraint	weight	target
STRETCH vowel	1.53	4.91
SQUEEZE vowel	0.00	1.47
SQUEEZE high	0.15	-1.86
SQUEEZE non-low	0.09	-21.60
SQUEEZE lax	0.13	-8.88
SQUEEZE closed	0.47	1.11
SQUEEZE complex coda	0.24	0.13
SQUEEZE pre-voiceless	1.67	0.76
SQUEEZE post m	0.04	-26.76
SQUEEZE unaccented	0.46	-0.43
SQUEEZE phrase-medial	0.92	0.17
Neg. Log Prob.	**3333.91**	
AIC	**6711.81**	

26. A shocker

- Some of the targets are incredibly negative; like −2.6 seconds!
- Lefkowitz has an explanation for this, which we will cover later on.

27. A hypothesis that worked better: the Consistent Variation Hypothesis

"When one category of sounds (or larger prosodic constituents), defined either by its phonological properties or by the context in which it occurs, shows more random, unconditioned variance in some phonetic variable than some other comparable category, it should also show more phonologically conditioned variation than that other category, appearing more susceptible to orthogonal phonological factors."

- This is a prediction of the *whole approach*, for high variation is the result of low constraint weight, irrespective of the source of variation (other constraints, knobs, randomness).
  - ➢ and that in turn, is the result of Harmony being additive
  - ➢ Take a look at our Lefkowitz's Theor em spreadsheet again to confirm.
- I.e. the prediction, unlike others, needs no auxiliary hypotheses to be made; all you need is MaxEnt.

28. Lefkowitz's method for testing the Consistent Variation Hypothesis

- Do this repeatedly:

- Imagine all the data sorted into "**bins**", each defined by setting every single experimental parameter (voicing of following consonant, nuclear accent, etc. etc.)
- Select a **grand factor**, like following voiced vs. following voiceless consonants.
- Each grand factor partitions the bins into two **bin-groups**.
- For phonologically conditioned variation:
  - take the mean of each bin in each bin-group
  - Then take standard deviation of these means, for each of the two bin-groups.
- For random variation:
  - Take standard deviation of each bin in each bin-group.
  - Average these across all the bins in each group.
- In brief:
  - Phonologically conditioned: s.d. of the means
  - Random: mean of the s.d.'s

29. Pursuing the method: the Grand Factors used for forming bin-groups

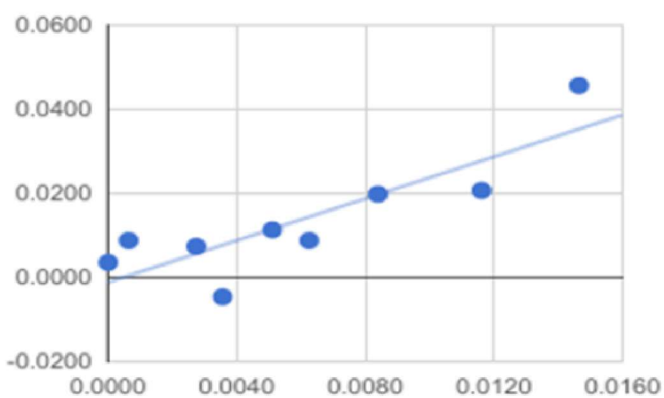
Bin-forming criterion	*Definition*
nasalonset	All the conditions with /b/ onsets except those with target words /bɛd/ or /bɛts/[32] were compared to all the conditions with /m/ onsets.
tense	All the closed-syllable conditions with /eɪ/ or /i/ except those with target words /meɪd/ or /meɪts/ were compared to all of the conditions with /ɛ/.
high	All the conditions with /eɪ/ or /ɛ/ were compared to all the conditions with /i/ or /ɪ/ except those with target words /mɪd/ or /mɪts/, since the corresponding mid vowel categories, involving words of the form /mɛd/ and /mɛts/, were not present in the stimuli.
low	All the conditions with /æ/ except for those with target words /mæd/ or /mæts/ were compared to all the conditions with /ɛ/ (conditions with /eɪ/ were excluded from this set of mid vowel data since there were no low tense vowels).
closed	All the open syllable conditions were compared to all of the conditions with /t/-codas and tense vowels (conditions with lax vowels were excluded from the closed syllable set because lax vowels do not occur in open syllables, and /d/ and /ts/ data were excluded to eliminate effects of voicing and coda complexity).
voiceless	All of the conditions with /d/ codas except those with target word /mɛt/ were compared to all of the conditions with /t/ codas.
complex	All of the conditions with /t/ codas except those with target word /mɛt/ were compared to all of the conditions with /ts/ codas.
accented	All of the accented conditions were compared to all of the unaccented conditions.
final	All of the phrase-final conditions were compared to all of the phrase-medial conditions.

30. Results of the method

- We are interested in comparing the *increase* in variation that we get when we look at the "favors longer" bin-group.

Bin-group	*Random*	*Phonologically Conditioned*
high (high vs. mid)	0.0027	0.0074
low (ɛ vs. æ)	0.0063	0.0088
tense (ɛ, ɪ vs. eɪ, i)	0.0000	0.0036
nasal onset (m vs. b)	0.0006	0.0088
closed (t vs. Ø)	0.0146	0.0457
voiceless coda (t vs. d)	0.0116	0.0207
complex coda (ts vs. t)	0.0036	-0.0046
accented	0.0051	0.0113
phrase-final	0.0084	0.0198

- Here is the scattergram; horizontal axis is random variation.



- $r = 0.86$, $p = 0.003$
- So things are looking good!
- Caveat: less clear when you work with log duration instead of duration.