

## Class 17, 5/25/2020: Lefkowitz II; Contrast in Phonology I

### 1. Assignments

- Homework #4 is due; hand in by email
- Read:
  - Flemming, Edward (2004) Contrast and Perceptual Distinctiveness. In Bruce Hayes, Robert Kirchner, and Donca Steriade, eds., *Phonetically-Based Phonology*, Cambridge University Press.
  - On course web site.
- Feel free to discuss progress/problems with your term project.
  - Office hours W, F at 2 and by appt.
- I can make appointments for your handout-session with me (deadline below)
  - Syllabus: “This should take place before the middle of Finals Week, Wed. 6/10, but ideally earlier. 5% of your grade.”

### MORE ON MAXENT PHONETICS

### 2. Lefkowitz so far

- Pursuing Harmonic Grammar phonetics along the lines of Flemming.
- Taking the MaxEnt variety of Harmonic Grammar to model variation.
- Theme: are there mathematical patterns that emerge from the math and are characteristically found in phonetic data?
- Very straightforwardly: Flemmingian parabolas generate Gaussian distributions.
- This is not true for duration, leading Lefkowitz to adopt hemiparabolic penalty functions.
- 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 checks out ok using the method of bins and bin-groups.

### 3. Detour: revisiting our Braver simulation

- Recall the goal:
  - Japanese [ki] is held to be bimoraic, but influenced by its base form [ki-ga].

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

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

- Last time: we modeled a distribution trained on the *mean* value of 54.99. This worked.
- ZL: more realistic to train on, and match, a distribution defined by the mean and standard deviation given (54.99, 21.89).
- This works fine, it seems, and yields a near perfect fit (spreadsheet).

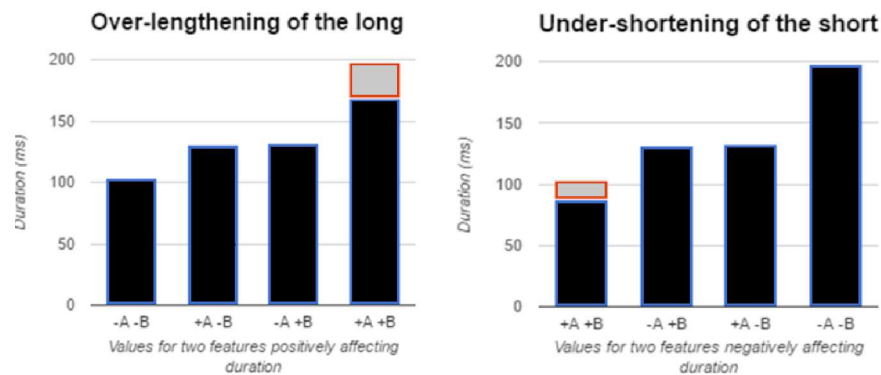
#### THE HYPERADDITIVE LENGTHENING GENERALIZATION

#### 4. As stated in Lefkowitz

“The duration of a segment taken to be undergoing both of two lengthening processes will be longer than the duration predicted by a multiplicative model, given the magnitude of the processes when they apply individually, and using the case where neither applies as the baseline. Equivalently, the duration of a segment taken to be undergoing both of two shortening processes will not be as short as predicted.”

- In sum:
  - More than multiplicative for lengthening effects, or equivalently,
  - Submultiplicative for shortening effects.

## 5. Lefkowitz's graph



## 6. Example

- The effects of pre-voiceless shortening are less for lax than tense vowels.
- Translate: the effects of non-pre-voiceless lengthening are greater for tense than for lax vowels.

## 7. Typology

- It emerges from his own empirical study of English vowel duration.
  - Tested by including an interaction term in the regression modeling, which tests significant.
- He cites a fair number of earlier studies, most of which also found it.

## 8. Theoretical basis

- Lefkowitz posits that most constraints are SQUEEZE constraints; i.e. don't shorten.
- His best grammar has just one STRETCH constraint (general, for all vowels), and a bunch of SQUEEZE constraints.

STRETCH vowel	1.57	3.19
SQUEEZE vowel	0.00	0.95
SQUEEZE high	0.04	-20.22
SQUEEZE non-low	0.32	2.86
SQUEEZE lax	0.02	-28.67
SQUEEZE closed	0.90	2.81
SQUEEZE complex coda	1.96	4.27
SQUEEZE pre-voiceless	1.56	0.75
SQUEEZE post m	0.03	-25.37
SQUEEZE unaccented	0.39	-0.34
SQUEEZE phrase-medial	0.77	0.22

## 9. Trying to replicate this in simple cases: spreadsheet

- I'm not quite there, perhaps you can help fix.

### SKEWNESS

## 10. Time for me to learn about skewness!

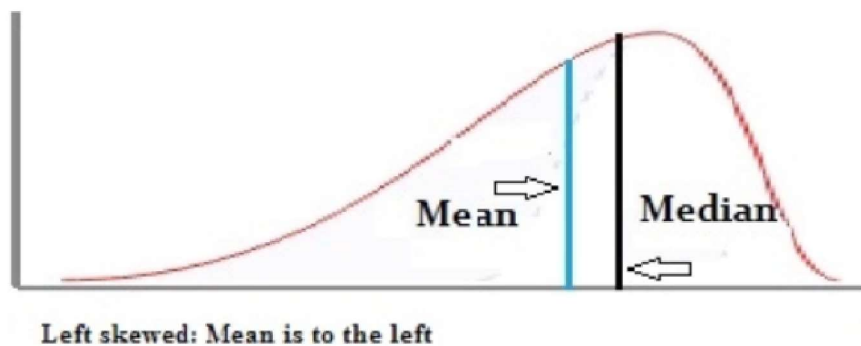
- I visited this page:  
➤ <https://www.statisticshowto.com/skewness/>
- I got this formula:

“The formula given in most textbooks is

**Skew = 3 \* (Mean – Median) / Standard Deviation.**

This is known as an **alternative Pearson Mode Skewness.**”

- They give this nice diagram:



- Due to big left-tail, mean is less than median, so
- **negative means steeper right side**
- **positive means steeper left side** (mirror image)

## 11. The behavior of Lefkowitz's model with respect to skewness

“whenever two categories of segment vary only by whether or not they are subject to some SQUEEZE constraint (for example, [1 +V] as compared to [2 +V], with respect to the constraint SQUEEZE[V/\_\_\_σ]), not only will segments subject to the constraint have shorter mean duration, their distribution should also have lower skewness, i.e. a steeper right side. STRETCH constraints are similarly able to result in higher skewness for the categories to which they apply.”

- We lack time to check this computationally.
- Lefkowitz verifies it empirically with his big experiment.

## 12. Summing up: is this theoretical phonetics?

- I see it as at least baby theoretical phonetics, since it has the relevant traits:
  - Precisely articulated theoretical framework.
  - General predictions about patterns, made with the math of the theory.
  - Checking the predictions against datasets.

# DISPERSION AND CONTRAST IN PHONOLOGY

## DISPERSION THEORY I: REVIEW OF 201A MATERIAL

## 13. References

- Flemming, Edward (2002) *Auditory representations in phonology*, Routledge.
- Flemming, Edward (2004) Contrast and Perceptual Distinctiveness. In Bruce Hayes, Robert Kirchner, and Donca Steriade, eds., *Phonetically-Based Phonology*, Cambridge University Press (readings)
- Early 2000's work of Jaye Padgett, UC Santa Cruz, all posted at <http://people.ucsc.edu/~padgett/papers.html>
- Current work of Juliet Stanton, NYU, all posted at <https://julietstanton.github.io/>

## 14. Dispersion

- The contrasting forms of a language should sound *different from each other*.
- This means that the “goodness” of a form depends on the other forms that are present.
- So the founders of OT were totally wrong in taking “harmonic completeness” to be an argument for the theory — it’s just not empirically true, where contrast effects are present.

## 15. The founders on harmonic completeness

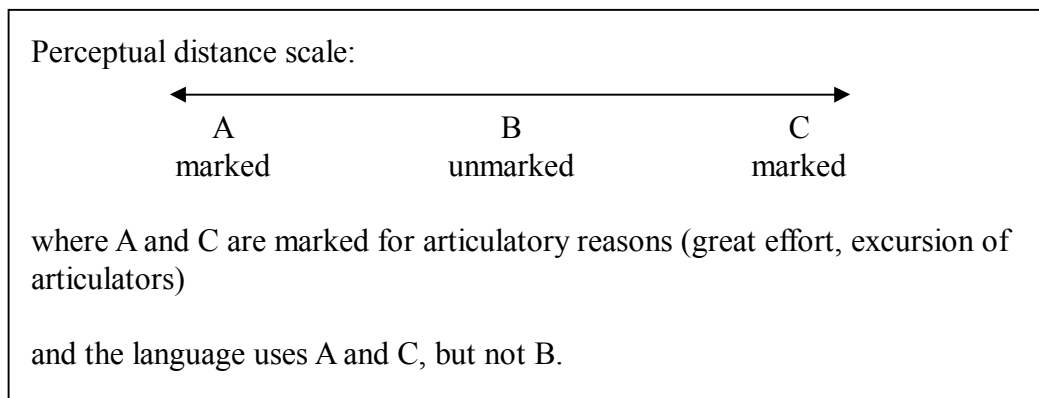
The harmonic completeness property (Prince and Smolensky 1993, ur-ms., §9.2.1) holds when:

If Segment A has a subset of the markedness violations of Segment B, then any inventory that includes B must include A.

or in Smolensky (2006, *The Harmonic Mind*, Ch. 14)

The segmental inventory of a language  $\mathcal{L}$  is **harmonically complete** with respect to some structural dimension  $d$  (e.g., place) if and only if the following holds: if  $x$  is a legal segment in  $\mathcal{L}$ , and  $y$  differs from  $x$  only in that it is more harmonic (less marked) with respect to  $d$ , then  $y$  is also a legal segment in  $\mathcal{L}$ .

## 16. Flemming's general argument: the "excluded center"



- This is problematic for OT, since if we simply ban the unmarked, the factorial-typology predictions are trashed.

## 17. Excluded center I: Russian Palatalization

- It is marked for consonants to be palatalized or velarized. But Russian has only palatalized and velarized consonants (Padgett 2001)<sup>1</sup>

## 18. Excluded center II: [Advanced tongue root] and breathy/creaky voice

- If laryngeal differences are used to phonetically reinforce an [Advanced Tongue Root] contrast, then every vowel can be marked: either breathy: [V̤] or creaky: [V̥].
- See Ladefoged and Maddieson, *The Sounds of the World's Languages* for background.

## 19. Excluded center III: vertical vowel systems (Flemming)

- Examples:
  - Marshallese
  - Kabardian and other Caucasian languages
  - Ndu languages (New Guinea)
- long vowel subsystem is normal; short vowel subsystem is "vertical":

i  
e  
a

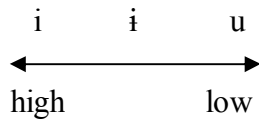
<sup>1</sup> Padgett, Jaye (2001) "Contrast dispersion and Russian palatalization," in Hume, Elizabeth and Keith Johnson, *The role of speech perception in phonology*.

## THE ANALYSIS OF THE EXCLUDED CENTER IN FLEMMING'S DISPERSION THEORY

### 20. The functionalist grounding of the approach

1. It is good to minimize articulatory effort.
  2. It is good to keep contrasting entities perceptually distinct.
  3. It is good to have contrasts, forming the basis of distinct words.
- Of course, these principles conflict with each other (☞ explain 1-2, 1-3, 2-3)

### 21. Analysis I: The F2 perceptual continuum



and similarly for various other continua

### 22. Analysis II: Constraints banning bad contrasts

MINDIST=F2:2 ‘Don’t allow contrasting entities to differ by less than 2 on the F2 scale’

- Let’s hold off for a moment on what these apply to ...

### 23. Analysis III: constraint requiring enough contrasts

MAXIMIZE CONTRASTS (perhaps along some particular dimension)

- MAXIMIZE CONTRASTS constraints award **check marks**, rather than penalizing with asterisks. One ✓ for each contrasting entity.

### 24. Analysis IV: Effort -Based Markedness Constraint

“\*HIGH EFFORT” cover term for a constraint that bans the production of peripheral vowels in short time frame. [i] and [u] are peripheral; [ɪ] is not.

### 25. The Factorial Typology of these Constraints

- a. **Vertical System:** \*HIGH EFFORT is on top (possibly tied with MINDIST).

	*HIGH EFFORT	MINDIST=F1:2	MAXIMIZE CONTRASTS
☞ [i]			✓
*[i]	*!		✓
*[u]	*!		✓
*[i-i]	*!	*	✓✓
*[i-u]	*!	*	✓✓
*[i-u]	*!*		✓✓
*[i-i-u]	*!*	**	✓✓✓

[☞ What candidates are harmonically bounded? Specify the bounders.]

### b. Polarized system

E.g. Spanish, with the peripheral vowels [i - u].

[☞ What is the ranking under which [i-u] wins? ]

- *This is claimed to be a case of the excluded center!*
  - But no one had realized this before, because this ranking is so common across languages.

### c. Rich system: Hopi, with [i i̯ u]

[☞ What is the ranking for this?]

DISPERSION THEORY IS A VERY DISTURBING DEVELOPMENT FOR OT

## 26. Contrast is a property of lexicons

Flemming (2004): “Constraints on the distinctiveness of contrasts evaluate relationships between forms. So if we want to determine whether a putative word is well-formed, we must consider whether it is sufficiently distinct from neighboring words. But these words must also be well-formed, which implies assessing their distinctiveness from neighboring words, and so on. Thus it seems that we cannot evaluate the well-formedness of a single word without determining the set of all possible words.”

## 27. Local environments as a means of structuring and containing the problem

“The analyses above avoid [the] problem [of mass comparisons] by considering only the evaluation of inventories of contrasting sounds (or short strings of sounds) in a particular context rather than evaluating complete words.”

- Thus the candidates are schematic bits of sound that form the parts of words.
- As with P-map theory, a clear notion of legal phonological contexts is needed to make the theory complete.



## 28. Meta-Analysis

“More generally, the strategy for avoiding the problem of mass comparisons is to derive generalizations about the set of possible words in a language – e.g. stressed vowels are all drawn from a certain set – rather than deriving particular words.”

- Hence, the set of legal forms need not be listed, but simply consists of all the forms that obey the generalizations derived by the theory.

## 29. History of this issue

- **Changing the theory** by replacing MAXIMIZE CONTRASTS with \*MERGE — more on this later.
- **Denial of the problem** (continuing to do OT, but not acknowledging that the theory is incompatible with OT).
- **Rejection of the approach** — don’t accomplish dispersion in the synchronic theory of phonology, but rather with diachronic (Boersma, Blevins, others).

## 30. Generalizing somewhat: three phenomena that appear to respond to dispersion

- The “**excluded center**” in phonotactics
  - This we have just seen: Effort is sacrificed to get good Distance and a decent number (2) of contrasting entities.
- **Enhancement** of outputs in alternation.
  - Probably the same
- The “**neutralization of despair**” in poor-contrast conditions.
  - “Neutralization of despair” is both static (phonotactics) and active (alternations).

## 31. An example of enhancement of outputs in alternation

- Vowel harmony is often thought to have a perceptual teleology (e.g., work of Abigail Kaun)
- Example:
  - ‘goat-plural’ and ‘sheep-plural’ in a non-vowel harmony language

to-pi ‘goats’  
tΛ-pi ‘sheep-pl.’

- in a vowel harmony language

to-pu ‘goats’  
tΛ-pi ‘sheep-pl.’

- The language sacrifices rounding contrasts in suffixes — a cheap price to pay.
- ☞ what are the costs and gains in Flemmingian terms? (see constraints in (25))

### 32. Examples of the “neutralization of despair”

- These are abundant and seem to occur wherever a contrast is made only in a subset of possible contexts.
- The work of Steriade in the 1990’s was key in pointing this out.
- Some neutralizations of despair, from Steriade:
  - obstruent voicing when a sonorant does not follow
  - retroflexion when a vowel does not precede
  - consonant vs. null when flanked by consonants (\*CCC vs. ✓CC)
- ☞ In Flemmingian terms, what constraints must be strong to induce a neutralization of despair?

### 33. Is there a fourth category? The Fortition of Rescue

- This would be: add major articulatory effort to rescue a contrast that is in danger in a particular context.
- I can think of some English examples:
  - /j/ → [j] / \_\_\_ i in *ye, yield, yeast*
  - similar fortition of /ɹ/ in *rural* ['ɹʌəl], *error* ['ɛɹ]
  - obligatory release (in normal speech) of prepausal /t/ after a stop: *act, apt* vs. *hack, nap*
- ☞ What are the Flemmingian priorities here?
- These cases perhaps often are in free variation with the neutralization-of-despair outcome:
  - I know speakers who don’t say the /j/ in *ye, yield, yeast*
  - In more casual style I can drop the /ɹ/ in *error*: ['ɛə]
  - Vernacular speech deletes, rather than releases, the /t/ in *act* etc.

### 34. The Diachronic alternative (Boersma and Hamann 2008)

- Boersma, Paul and Silke Hamann (2008) The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology* 25: 217-270
- Dispersion is the natural response of a system that optimizes perceptual accuracy by learning from errors.

Roughly:

- You are a child, it is bedtime in a religious family, and you hear a parent say:  
[ˈwɪljəm, ˈsiː jə ˈpreɪz]
- You reason, “Uh-oh, that clearly was /seɪ/.” Better tighten (i.e. lower) my F1 criterion for /i/.”
- Such adjustments will do a sort of back-and-forth for “medial” phonemes like mid vowels.
- But for peripheral items, there are no “other side” errors to recentralize the phoneme.

- The same would work even if the child's percept were "20% probability for /i/, 80% for /eɪ/—the remorseful inner comment is based on a danger that could have been avoided.
- Boersma and Hamanns flesh out this scheme with a concrete learning simulation using Stochastic OT and the Gradual Learning Algorithm.