Linguistics 201A
Phonological Theory I

Winter 2023
B. Hayes

Class 15, 2/28/23: Phonetics in phonology III

1. Bureaucratic

- I need to move my office hours to 4 p.m. today (Christine Prechtel Ph.D. defense).
- Please make an appointment with me if this time is not good for you.

2. Current assignments

- Please hand in your Indonesian homeworks (extensions have been/can be granted).
- BH return and discuss the MaxEnt phonotactics homeworks.
- The UCLA course Web site is still down: please visit instead:
  - https://www.palisadessymphony.org/temp/index.htm
- Read for next time:
  - Braver, Aaron (2019) Modelling incomplete neutralisation with weighted phonetic constraints. *Phonology* 36:1-36.
  - Posted on the site listed above
  - No summary required

BRIEF COMMENTS ON THE MAXENT HOMEWORK

3. What are my feelings?

- The class rose to the occasion, learning the skills quickly; bravi.

4. The topics covered

- Lardil vowels, Italian onsets, Swahili vowels, Guarani vowels

5. Good-practice hints that emerged from my reading

- A list of constraints with fitted weights probably always helpful; this always appears in journal articles, I believe.
- Make your scattergram have axes with equal scales and lengths.
- Likelihood ratio test reminder: compute difference in likelihood resulting from removing a constraint, double it, then apply =CHIDIST(*d*, 1) to the difference *d*.
- A couple of homeworks did something nice, if space-consuming:
  - a series of scattergrams and likelihood values as model improves; sheep becomes snake.

6. Parasitic harmony

- Term from an old paper by Steriade, perhaps receding in use

- Cases from the homeworks:
  - AGREE(round) if high (Lardil)
  - AGREE(back) if [−low] (Swahili)
- These cases are entirely characteristic and are the subject of a whole literature in the “Agreement by Correspondence” (ABC) literature; most cited article is:
  - Rose, Sharon, and Rachel Walker. "A typology of consonant agreement as correspondence." *Language* (2004): 475-531.

STARTING UP AGAIN

7. Phonetics in phonology: so far

- Basics of perceptual cues: external and internal
- The P-map as a basis for default constraint weightings (Steriade, Zuraw)
- Deriving cross-linguistic asymmetries in the direction of assimilation
- A gradient solution to the Too Many Solutions problem.

8. Phonetics in phonology: plans for remaining class(es)

- Testing the P-map experimentally and computationally (White)
- The P-map and the Dilemma of Dispersion
- Generative phonetics I: rule-based theory and the English berry sentences
- Generative phonetics II: Harmonic Grammar and near-neutralization

WHITE ON SALTATION

9. References

- Typology:
  - Hayes, B. & White, J. (2015). Saltation and the P-map. *Phonology*, 32(2), 1–36.
- Experiments on adults:
  - White, J. (2014). Evidence for a learning bias against saltatory phonological alternations. *Cognition*, 130(1), 96–115.
- On infants:
  - White, J. & Sundara, M. (2014). Biased generalization of newly learned phonological alternations by 12-month-old infants. *Cognition*, 133(1), 85–90.
- MaxEnt modeling:
  - White, J. (2017). Accounting for the learnability of saltation in phonological theory: A maximum entropy model with a P-map bias. *Language*, 93(1), 1–36.

10. Ur-reference, inventing the methods White used

- Wilson, Colin (2006) Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science* 30:945-982.

11. Basic scheme

- This is an Artificial Grammar Learning experiment, modeled on Campadanian — is the saltatory pattern hard to learn?
- Stimuli (singular-plural of pictures)

p	b
f	v

A. ap avi
B. ab abi

- Participants get trained up, then reponds with plurals to both training forms and novel "wug-of-wug" forms.

12. Key results

- Participants tend to generalize [ap] ~ [av-i] to [ab] ~ [av-i], when not exposed to [ab].
- They do this to some extent even when trained explicitly on [ab] ~ [ab-i].
- A followup on 12-month-old infants yielded similar results.

13. You can train a MaxEnt grammar to be a White-participant

- The following tableaux simply mimic the experimental responses.

(11) Calculating predicted probabilities in tableaux

a. Input /VpV/ in experiment 1, potentially saltatory condition

/VpV/	*V[−voice]V 2.20	*MAP(p, v) 2.17	*V[−cont]V 1.86	*MAP(b, v) 1.30	Penalty score	$e^{(-\text{penalty})}$	Predicted prob.
VvV		1			2.17	.1142	.87
VpV	1		1		4.06	.0172	.13

b. Input /VbV/ in experiment 1, potentially saltatory condition

/VbV/	*V[−voice]V 2.20	*MAP(p, v) 2.17	*V[−cont]V 1.86	*MAP(b, v) 1.30	Penalty score	$e^{(-\text{penalty})}$	Predicted prob.
VvV				1	1.30	.2725	.64
VbV			1		1.86	.1557	.36

14. Three levels of modeling

- **Flat-out data matching**, from the training data (or, if you're working with a real language, a corpus).
- **Exploratory modeling**, seeing what grammar would reflect the experimental responses.
- **Try to mediate**: UG-biased learning from the training data, trying to mimic the response data.

15. The effort made at biased modeling

- Obtain an approximation of the P-map: take a confusion matrix, and set up an artificial MaxEnt grammar: "If I hear [*x*], what is probability that I perceive [*y*]?"
- Here is one confusion matrix that White employed, from Cutler et al. (2004, *JASA*):

TABLE V. Confusion matrix for initial consonants at 0 dB SNR categorized by the Dutch listeners. Percentages of correct responses have been pooled over participants and vowel contexts.

	Response																					
	pie	tie	car	far	thin	see	she	chin	hi	be	do	go	very	there	zoo	joke	yell	my	no	lie	row	win
Stimulus	p	t	k	f	θ	s	ʃ	tʃ	h	b	d	g	v	ð	z	dʒ	j	m	n	l	r	w
p	**30.8**	3.3	9.2	9.6	2.9			0.4	19.2	11.7	1.3	1.3	2.9	1.7	0.4		0.8	0.8	1.3	1.3		1.3
t	24.6	**14.2**	12.5	7.5	7.9	0.8		2.9	11.3	7.1	0.4	2.1	1.3	1.7				2.1	3.3	0.4		
k	25.0	7.9	**25.8**	3.8	4.2	0.4	0.8	0.4	13.8	4.2	1.3	3.8	1.3	0.8	0.4	0.4	1.3	0.4	1.7	1.3	0.4	0.8
f	24.6	2.1	9.2	**15.0**	7.1		0.4	0.4	9.2	15.0	1.7	2.9	5.4	4.2		0.4	0.4	0.4	0.4		0.4	0.8
θ	18.8	6.3	3.8	13.3	**12.1**	0.4	0.4	0.4	7.1	14.2	2.5	1.7	2.9	7.5			0.4	1.3	2.9	2.9		0.8
s	0.4	2.5	0.4	12.5	24.6	**30.4**	0.8	0.4		0.8	1.3		3.3	7.9	14.6							
ʃ		0.4			1.3	6.7	**72.5**	18.3									0.8					
tʃ	3.3	4.2	1.3	2.1	2.5	1.3	4.6	**70.8**	1.3	1.3	0.4		0.4	0.8		5.4	0.4					
h	26.3	4.6	12.1	11.3	5.0	0.4	0.4	0.8	**17.9**	8.3	1.3	0.4	4.6	1.7	0.4		0.8		0.8	1.7	0.8	0.4
b	7.5	0.4	5.8	9.2	1.7	0.4		0.4	12.5	**28.3**	2.5	0.4	4.6	2.1			2.9	7.1	2.9	5.0	1.3	5.0
d	2.5	2.1	1.3	1.3	5.4				8.8	12.1	**10.8**	2.5	2.1	12.9	0.4	0.4	6.3	4.2	12.5	12.5		1.7
g	3.3	1.3	9.2	2.9	2.5		0.4	1.3	9.2	10.0	5.0	**17.1**	1.7	3.3		0.8	24.2	0.8	2.5	2.5	0.4	1.7
v	7.5	2.9	2.5	8.8	5.0	0.4			6.7	30.0	1.3	1.7	**9.6**	7.9			2.1	3.8	1.7	0.4	2.5	5.4
ð	2.5	1.3	2.5	1.7	14.6	2.1	0.4	0.8	2.1	17.1	10.0		1.3	**18.8**	1.3	1.7	1.3	1.3	3.3	12.1	0.8	2.9
z		0.8	1.3	1.3	9.6	3.3	0.4	1.3		7.9	5.0		2.5	23.8	**27.1**	1.3	2.5	2.1	5.0	0.4	0.8	3.8
dʒ	4.2	0.4	2.5	0.4	2.1		0.8	18.3	2.1	2.1	6.7	2.1		7.5		**40.4**	5.8	0.4	1.3	2.9		
j	1.3		0.8	0.8	0.8		0.8	0.4	2.1	4.2	2.5	1.3	0.4	1.3	0.4	4.6	**69.6**	2.5	4.2	1.3		0.8
m	3.8	0.8	2.5	2.9	0.4				2.1	9.6	0.8		2.5					**50.0**	10.0	5.0	5.4	4.2
n					0.4				2.1	1.7	1.3					0.8	0.4	12.9	**73.8**	4.6	0.4	1.7
l	5.8	1.3	1.7	1.3	1.3			0.4	2.1	8.3	1.7	0.8	2.5	3.3			1.3	10.0	4.2	**46.7**	2.1	5.4
r	2.5	1.3	1.7	2.1				0.4	5.8	14.6	0.8	2.1	1.3	0.8				0.8	0.4	0.4	**58.3**	6.7
w	1.7		0.4	0.4	0.4	0.4			1.7	5.8	0.8		2.1	0.4				5.8	0.8	2.5	1.7	**75.0**

J. Acoust. Soc. Am., Vol. 116, No. 6, December 2004 Cutler *et al.*: Native and non-native phoneme confusion 3673

- The artificial MaxEnt grammar gives the default weights for *MAP() constraints.

16. White's method for biased learning

- Taken from Wilson (2006), with improvements.
- Here is the MaxEnt formula from Goldwater and Johnson (2003):

$$\log \mathrm{PL}_{\bar{w}}(\bar{y}|\bar{x}) - \sum_{i=1}^{m} \frac{(w_i - \mu_i)^2}{2\sigma_i^2}$$

  - ➢ First part: log likelihood of the batch of candidates *y* given their URs *x*.
  - ➢ Second part: the prior, which implements the bias

17. Key part of the prior

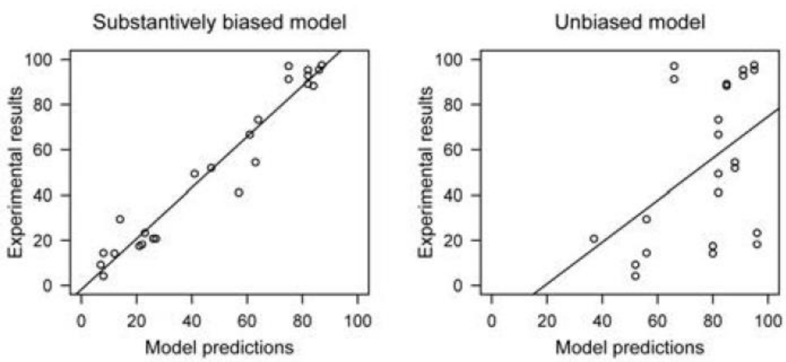
- Find the difference between the weight being solved for, and its ideal values, μ.

- Square it
- Rescale it
  - You could just multiply it by something, but instead one divides by twice sigma squared. I will briefly tell you why.

18. The point of doing this

- The grammar learned emerges as a compromise between matching the learning data and matching "UR" in the form of the prior.
- Socrates:
  - If σ is enormous, what do you match?
  - If σ is tiny, what do you match?

19. The method worked well for White



- It would appear that the participants actually didn't pick up much from the actual training data.
- Their respondes were heavily influenced by the similarity prior.
- Real-life learners of Campidanian had a far better chance to overcome their prior.

THE P-MAP AND THE DILEMMA OF DISPERSION

20. Some 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)
- Work of Jaye Padgett, UC Santa Cruz, a guide posted at https://humweb.ucsc.edu/jayepadgett/wp/research
- Work of Juliet Stanton, NYU, posted at https://julietstanton.github.io/
- Work of Paul Boersma and colleagues, noted below

21. Dispersion

- The contrasting forms of a language should sound *different from each other*.
- This is completely sensible a priori, given the demands of speech perception.
- It means that the "goodness" of a form depends on the other forms that are present in the language.
- This ends up falsifying the hypothesis of "harmonic completeness," as defined below.

22. Harmonic completeness

- The **harmonic completeness** property (Prince and Smolensky 1993, §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.
- It's very tricky to deduce harmonic completeness from OT (depends on Faithfulness constraints) — can't cover it here.
- Flemming: harmonic completeness as an *empirical prediction* is very far from the truth.
  - ➢ Not due to random weird cases, perhaps with a historical origin.
  - ➢ Rather, false on a principled basis; we should *expect* it to be false.

23. Steriade applied the P-map to dispersion

- Straightforwardly: Rich Base, then phonological neutralization of contrasts that are too close.
- Surviving sounds are dispersed.
- Not clear this will work … because of Flemming's discovery

24. Flemming's general argument: the "excluded center"[1]

Perceptual distance scale:

◄──────────────────────────────►

A	B	C
marked	unmarked	marked

where A and C are marked for articulatory reasons (greater effort, excursion of articulators)

and the language uses A and C, but not B.

- This is problematic for OT!

[1] This helpful term due to Boersma and Hamanns, discussed later.

- Suppose we adopt a constraint banning B — this would be very bad, because: B is what occurs when there is no contrast on the dimension.

25. 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).[2]
  - ➢ And many languages like English or Spanish use non-palatalized, non-velarized consonants.

26. Excluded center II: choice of a high vowel row

- [ɨ] normally occurs in a high vowel row only where [i] and [u] are already present.

27. Excluded-center cases are ordinary — what theoretical account could predict their normality?

- Synchronic — within phonological theory: Flemming and his successors.
- Diachronic — via a theory of phonetic acquisition: Boersma and Hamann, others

FLEMMING'S ANALYSIS OF THE EXCLUDED CENTER IN BRIEF

28. Flemming's dispersion theory

- Flemming, Edward (2004) Contrast and perceptual distinctiveness. In Hayes, Kirchner, and Steriade, *Phonetically based phonology*. Cambridge: Cambridge University Press.

29. The functionalist basis of the constraints

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.

- These principles conflict with each other
  - ➢ ☞ explain 1-2, 1-3, 2-3

30. Analysis I: The F2 perceptual continuum

i ɨ u

⟵⟶

high low

and similarly for various other continua

[2] Padgett, Jaye (2001) "Constrast dispersion and Russian palatalization," in Hume, Elizabeth and Keith Johnson, *The role of speech perception in phonology*.

31. Analysis II: Constraints banning bad contrasts

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

- Violated when the inventory includes both [i] and [ɨ], or [ɨ] and [u].
- Let's hold off for a moment on what representations MINDIST constraints apply to …

32. 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.

33. 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; [ɨ] is not.

34. 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
☞ [ɨ]			✓
*[i]	*!		✓
*[u]	*!		✓
*[i-ɨ]	*!	*	✓✓
*[ɨ-u]	*!	*	✓✓
*[i-u]	*!*		✓✓
*[i- ɨ -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.
  - ➢ There really are "vertical vowel" systems, in which /ɨ/ is the only high vowel (albeit with lots of allophony).

c. **Rich system**: Hopi, with [i ɨ u]

[☞ What is the ranking for this?]

PROBLEMS WITH DISPERSION THEORY

35. What is a theory of grammar for phonology/phonetics?

- We think of it as a production system: input strings of lexical UR's from the morphology/syntax and output phonetic form.
- We can also invert the system, perhaps with Bayes's Theorem, to use it in perception.

36. Contrast is a property not of derivations but 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."

- Flemming is quite candid that his system is a system of "language design", not phonology.

37. Flemming-grammars *can* be implemented … sort of

- A squib coming out in *Linguistic Inquiry* (author unknown, I am a reviewer) finds a way to evaluate candidates in Flemming's system.
- The input is the entire set of URs, and the candidate are all combinations of candidates for the entire set of UR's.
- This is somewhat discouraging in practical terms, since, conservatively, there must about 20^{10000} candidates.
- But at least this is a rigorous formalization, which might not be true of all work that has pursued Flemming's ideas.

38. A 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.
- There exists a mechanism that could produce children who prefer contrast.

39. The "detectible error" scenario

Example:

- You are an infant with older siblings *Lou* [lu] and *Lee* [li]. You know their names.
- Your parent yells at either Lou or Lee from time to time, inadvertently helping you to tune your /i/-/u/ distinction.
- Your parent yells at Lou, and you mistakenly hear them say
  [ˈkʌɾ ɨɾ ˈaʊt, li] !!!
- You reason, "Uh-oh, that clearly was [lu]." I'm too predisposed to hear [i]! I'd better tighten (i.e. raise) my F2 criterion for [i]."
- Similar episodes (mishearing [u] for intended [i]) lead the infant to lower her F2 criterion for [u].
- Hence, the perceptual standard for this vowel becomes more stringent — more dispersed.

GENERATIVE PHONETICS

40. Usage

- Caution: the use of this term is **entirely non-standard**, and reflects a point of view.

41. One possible answer

- A **generative phonetics** would be the portion of a generative grammar that models the phonetic capacities of people.
- It would take the form of a formalized grammar.
- Like other grammars, it is intend to make predictions about future-gathered data.

42. What a phonetic grammar should predict

- Given a surface phonological representation and other factors (like speaking style, speaking rate, word frequency), what is the contour that the speaker will create for:
  - F0
  - formants
  - tongue body coordinates
  - … and durational pattern for all of the above
- I.e. generate a "movie of the mouth", or a waveform.
- The grammar is the algorithm for speaking

43. Four influential early works from the rule-based era in generative phonetics

- *From text to speech; The MITalk system* (1987) Jonathan Allen, M. Sharon Hunnicutt and Dennis Klatt (with Robert C. Armstrong and David Pisoni): Cambridge University Press, Cambridge.
- Pierrehumbert, Janet (1980) The phonology and phonetics of English intonation, MIT diss.

- Liberman, Mark and Janet Pierrehumbert (1984) Intonational invariance under changes in pitch range and length. In Aronoff and Oehrle, *Language sound structure*, MIT Press.
- Pierrehumbert, Janet, and Mary Beckman (1988) *Japanese tone structure*, MIT Press.

44. Key idea in this work

- The structural elements of surface phonology can be translated into *targets* defined in time and space (both physical and acoustic).
- The quantitative values assigned to targets obey simple arithmetical regularities.

45. Example: the Liberman-Pierrehumbert "berry" sentences

- The authors and colleagues at Bell Labs recite sentences consisting of lists of berries, using downstepping intonation: "Blueberries, bayberries, raspberries, mulberries, and brambleberries"

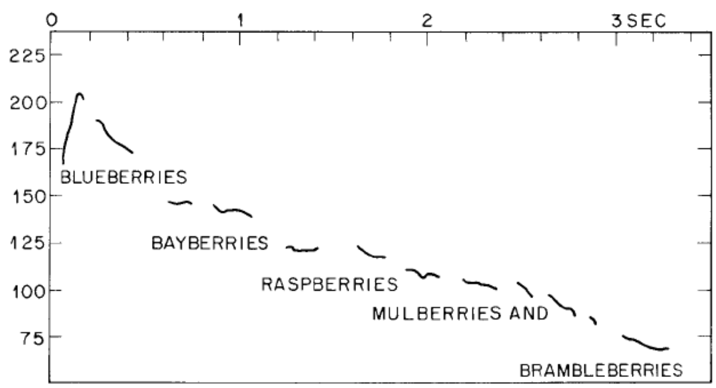


Figure 13
An F0 contour for the berry list *Blueberries, bayberries, raspberries, mulberries, and brambleberries*, produced with a sequence of step accents. Each step is smaller than the one before, so that the step levels appear to trace out an exponential decay.

46. Full dataset for one speaker

- The lists may consist of 2-5 berries.
- They did three pitch ranges (of which we will consider just one).

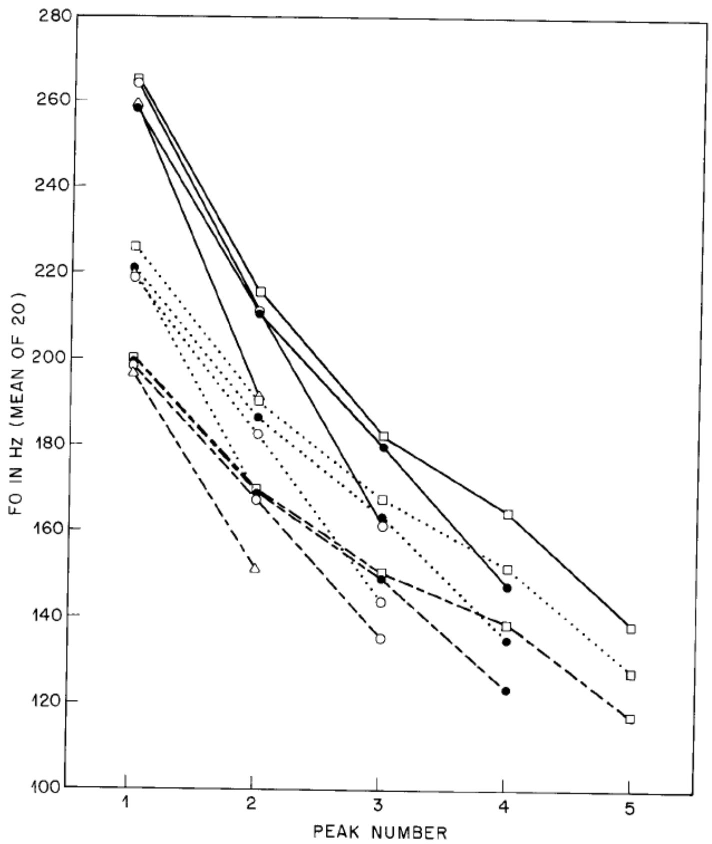


Figure 19
Downstep data (3 pitch ranges, 4 lengths) for subject DWS

47. The key ideas of the analysis

- There is a bottom line, as low as the speaker is willing to go.
- For any degree of emphasis, you can define a pitch scale.
- On this scale, each pitch accent is a constant multiple of the preceding one.
- … except that the last accent gets further multiplied by a Final Lowering constant.

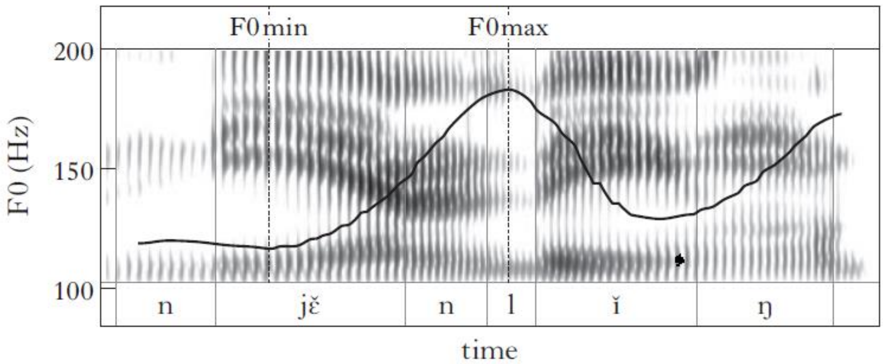
48. What was special about the early work on F0?

- It mostly involved *uncrowded targets* — each target is actually achieved.

49. What is needed for crowded targets?

- Example: a Mandarin rising tone after a high. I've drawn a dot where the L target might be.

L H L H



Pitch tracks and spectrogram of [njɛ̌nlǐŋ] 'age'.

- Crowded targets are compromises, usually between effort (rapid motion) and perceptability (achieve the target as much as possible).
- This led Flemming (2001, *Phonology*) to conceive of Harmony-based phonetic grammars:
  - Minimize a summed penalty of constraints penalizing effort and target-deviation.
  - Violations are gradient and based on the physical/psychoacoustic realm.