Linguistics 219 Spring 2018
Phonological Theory III B. Hayes

Class 18, 6/5/2018: Generative Phonetics I

1. Assignments

- Last reading:
  - I'd like to assign Flemming and Cho (2017) (harmonic grammar phonetics)
  - If you've just read it and would like Flemming (2001) instead we could read it; decide in class.
- Make your appointment to give a talk to me, with handout.

BRIEF COMMENTS ON THE HOMEWORK

2. The choice of base surprised me

- The most obvious choice, from historical perspective, is the local base:

4 syllables	[bìjaksána]
5 syllables	[[bìjaksaná] an]
6 syllables	[[[bìjaksàna] án] ña]

Preserving the stress of the intermediate base, you would get *[[bìjaksana] án] ña], which is phonologically legal (cf. [[kòntinuasí] ña]).

Note that English uses the local base, not the stem: [[[origin] ál] ity].

- Research such as Steriade's makes the choice of the stem as a base seem much more reasonable than it would have seems when Cohn did this work.

3. What is inherited?

- If I haven't erred, there is only evidence for inheritance of *stresslessness*; when I split up Faith-OO-Stress into two parts, the part requiring inheritance of stress proved unnecessary (OTSoft).

WHAT IS GENERATIVE PHONETICS?

4. Usage

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

5. What would be a generative phonetics?

- It 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 would make predictions; i.e. …
- Given a surface phonological representation and other factors (like speaking style, speaking rate, word frequency, i.e. our knobs), 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 an acoustic representation affiliated with it.
- … the algorithm for speaking

6. Three influential works in generative phonetics

- Pierrehumbert, Janet (1980) The phonology and phonetics of English intonation, MIT diss.
- Liberman and Pierrehumbert (1984) Intonational invariance under changes in pitch range and length. Readings.
- Pierrehumbert, Janet, and Mary Beckman (1988) *Japanese tone structure*, MIT Press.

7. Methodology I: as generative grammar

- Generative grammars can be altered and improved to achieve better fit to data.
- So can the assumptions about grammar content, theoretical architecture.
- The range of data explored can gradually be expanded.
- An article of faith of generative linguistics is that patient and intelligent work on these lines will gradually pay off.
- For articulate remarks along these lines, see Pierrehumbert and Beckman (1981: ch. 1).

8. Methodology II: as experimental science

- It is impossible to assess a phonetic grammar without measurement data.

9. Difficulty in generative phonetics

- **Too many grammars:** Like every branch of generative grammar, it suffers from the fact that there are so many conceivable ways to go about the task — how can we make informed choices?

- **Not even an single decent grammar:** But far more than in phonology, it is extremely hard to find any sort of grammar that works well for a non-trivial set of data.[1]
  - It is startling how bad the speech synthesis of extremely well-funded industrial labs is — *they* clearly have not yet solved the problem of generative phonetics!

10. Some research traditions in or related to generative phonetics

- Articulatory phonology (e.g. http://www.haskins.yale.edu/research/gestural.html, and a vast literature.)
- The ongoing research program of Paul Boersma (http://www.fon.hum.uva.nl/paul/).
- Work by mainstream phoneticians sometime create generative models. One that I have enjoyed is:
  - Andreas Windmann, Juraj Šimko, and Petra Wagner (2015) Optimization-based modeling of speech timing. *Speech Communication* 74:76–92.
- I suspect that there is tons of material I haven't read, and the presentation here is necessarily sketchy and suggestive.

11. The importance of tone and intonation

- Pierrehumbert and Beckman emphasize:
  - Mappings from phonological values to articulatory trajectories/formant trajectories are appallingly complicated for vowel and consonant features.[2]
  - I.e. nonlinear, many-to-one/one-to-many for features and measurable values.
- F0 is "sort of" clean by comparison, suffering from
  - modest vowel-height effects
  - consonant perturbations
  - voiceless cutouts
  - difficulty of measurement under creak
  - need for even more speaker normalization than is usual
  - need for expressive range normalization

---

[1] In this respect I think phonetics resembles syntax, where the extreme difficulty of the material means that accurate large-scale grammars are not widely pursued.

[2] The lab of Bryan Gick at UBC is a leading center for untangling all this; Gick thinks it may be simpler than this once you have the right theory of movement …

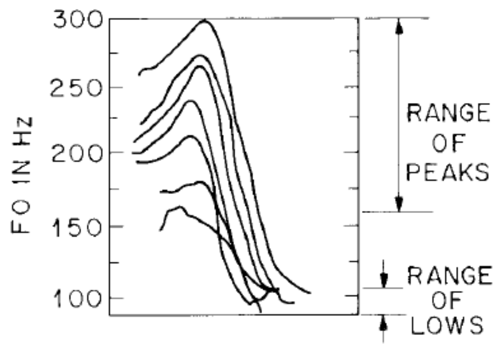


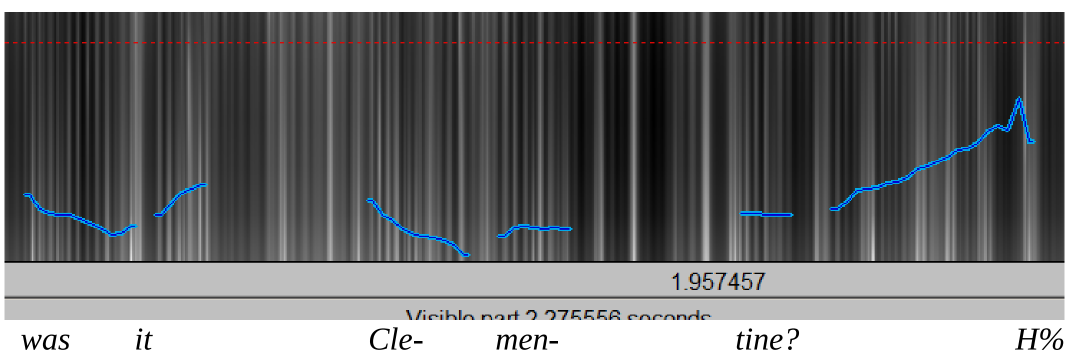
Figure 4
A declarative intonation pattern produced on *Anne* with seven different amounts of overall emphasis.

12. The concept of target

- The approach taken by Pierrehumbert and colleagues is sometimes called "targets-and-interpolation" theory.
- Phonological entities — often autosegments — translate as targets with physical values.

13. Alignment of targets in time

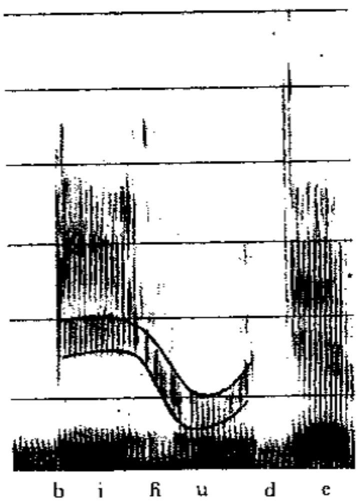
- This traditional advanced beyond traditional autosegmentalism in letting targets actually occur at boundaries.
- Thus in English, if we have a syllable that is post-nuclear with rising question intonation, then the H% tone is likely to create a rise to the very end:



INTERPOLATION

14. Interpolation and surface underspecification

- This is how you get phonetic values on regions having no target.
- See P. A. Keating: "Underspecification in phonetics", Phonology 5.2, 275-292 (1988), who offers this spectrogram of a Persian speaker saying [bihude].

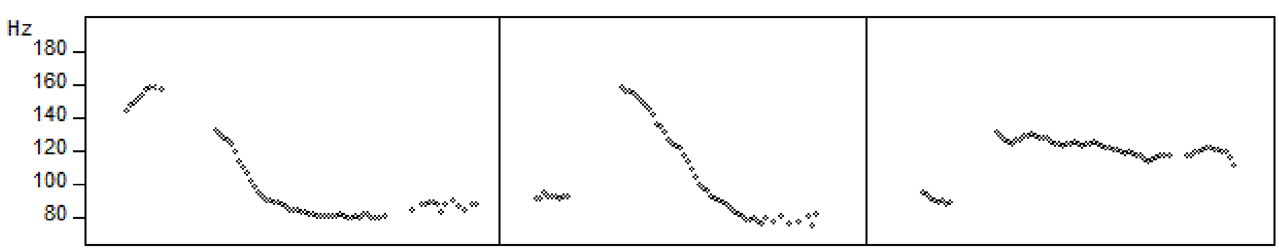


15. Kinds of interpolation

- The Liberman/Pierrehumbert paper, as well as *Japanese Tone Structure*, use basically simple linear interpolation.
- Augmented (for synthesis) with:
  - Smoothing (forward only)
  - insertion of random noise (increasing auditory realism)[3]
- I have seen other forms of interpolation
  - Slight, constrained, droop ((1991) Bruce Hayes and Aditi Lahiri "Bengali intonational phonology". *Natural Language and Linguistic Theory* 9: 47-96.). The droop never goes below the horizontal; i.e. one endpoint must be a minimum.
  - I've seen Japanese pitch tracks that look "billowy" to me.

16. A bit of Japanese phonology we will need

Japanese pitch accent



háʃi-o nuɾu	haʃí-o nuɾu	haʃi-o nuɾu
chopsticks-acc. paint	bridge-acc. paint	end-acc. paint
'paint chopsticks'	'paint a bridge'	'paint the end'

[3] I wonder if real noise is caused, not random, and might sound better?

- These illustrate a three-way contrast in disyllables of accent location/existence.
- In the olden days:

```
háʃi-o nuɾu            haʃí-o nuɾu         haʃi-o nuɾu
 \  \/                  |  | \/              |  \/
  H  L                  L  H  L              L  H
```

- In the view initiated by Pierrehumbert and Beckman:
  - ➢ Accent is phonemic, and is represented as a H*+L tone (prominent H, immediately falling L), H* linked to the phonemically relevant mora).
  - ➢ Words group into little units called accentual phrases.
  - ➢ There is a final L% boundary tone.

```
[ háʃi-o nuɾu ]            [ haʃí-o nuɾu ]       [ haʃi-o nuɾu ]
 |  |          |            |/  |        |        |/   |        |
L  H*+L       L%           L   H*+L     L%       L     H        L%
```

  - ➢ H*+L is a rapid and extensive fall.
  - ➢ Initial L is quite a bit lower when it can dock onto a mora.
  - ➢ Final L% is greatly lowered by a preceding H*+L

17. Proof of interpolation

- The required experiment demands that you systematically change that distance (in moras, syllables, or whatever) between a H and a L, then show that the slope changes in response.
- Why? because there can be rival explanations, notably *declination* (a purely mechanical downward drift in tone).
- Here is Pierrehumbert and Beckman's proof of interpolation between H and L in Japanese.
- Explaining the phonology
  - ➢ Glossing:

moriya-no	mawari-no	oˈmawarisan
proper name-of	neighborhood-of	policeman

  - ➢ accentual and higher phrasing:

    [[moriya-no mawari-no] [oˈmawarisan]]

  - ➢ tones with their ur-affiliations:

```
[ [ moriya-no   mawari-no ] [ oˈmawarisan ] ]
| |                         |     |          |
LH                          L     H*+L       L%
```

- tones once docking rules have applied:

```
[ [ moriya-no   mawari-no ] [ oˈmawarisan ] ]
      | |                   |/    |  \       |
      LH                    L     H*+L       L%
```

- Others: substitute shorter or longer words for *moriya*

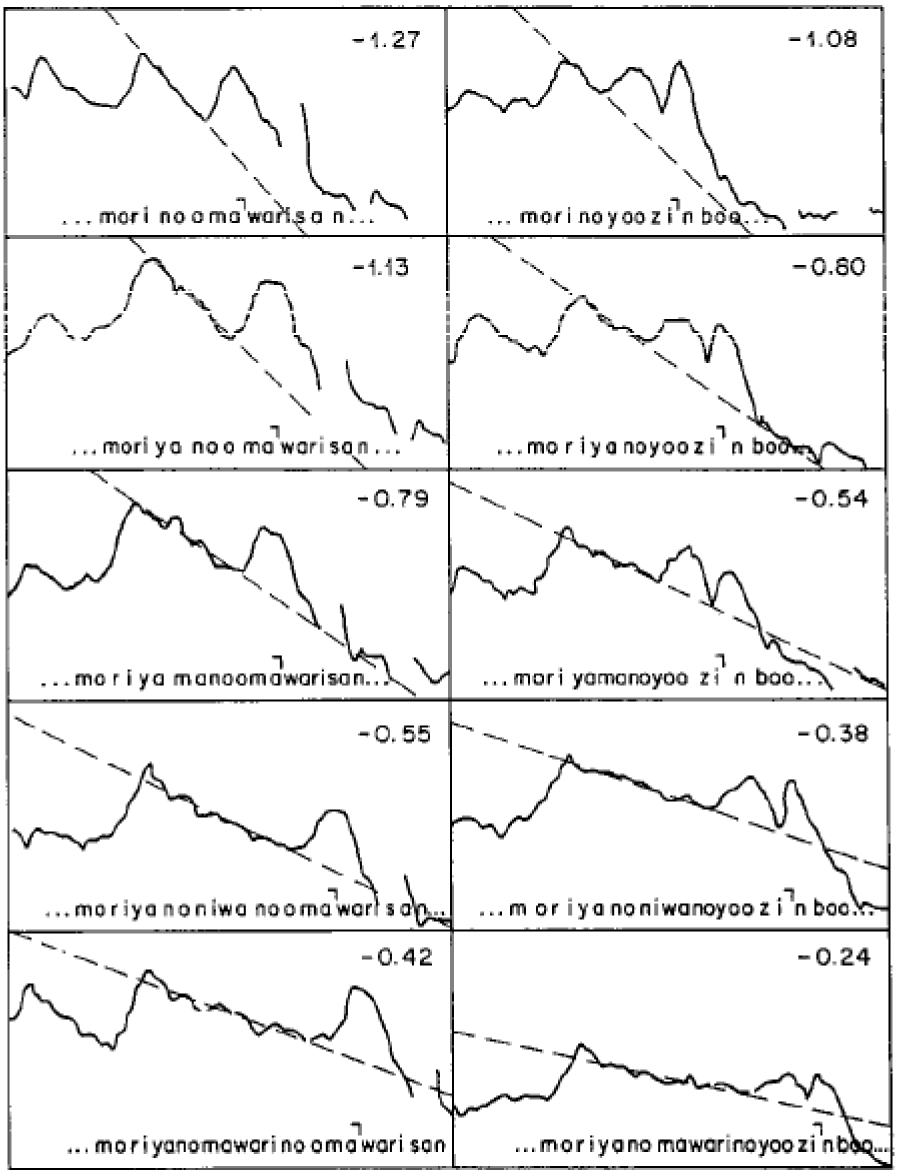


Figure 1.3

Unaccented phrases with 3, 4, 5, 7, and 8 morae before accentual-phrase boundaries with two different allophones of L%. The dashed line in each panel is a regression line fitted to all f_0 values between the peak for the phrasal H in the first phrase and the minimum for the interphrasal L%. The number in the upper right-hand corner is the slope of the regression line. Here and in subsequent figures a right corner in the transcription indicates the location of an accent.

18. More general comments on experimentation in generative phonetics

- *Phonological* rigor is crucial.
- I.e. we seek *minimal pairs*, or as close as possible to it given the structure of the language.
- Minimal pairs are the phonological side of sound experimental methodology in general, where we seek causes and effects by varying one thing at a time.

ALLOPHONY AND NEAR-NEUTRALIZATION

19. Traditional wisdom

- Without instruments, people only heard the L on words beginning with CVCV… where first CV is unaccented.

20. Closer to the truth, with measurement

- The L is *always* there.
- If there is an initial light syllable onto which it can hop, it does so (see above).
- Else it sits on the left boundary and gets a weak target.

21. More near neutralization

- The boundary H of Japanese is just a little bit lower than the accent H — something never noticed before experiments were done.

CLOSELY VS. DISTANTLY SPACED TARGETS

22. A cozy domain to investigate

- The success of the 1980's Pierrehumbertian experiments in finding orderly patterns suggests that the targets were characteristically *achieved* — there being enough room to do it.
- The one case of tonal crowding is suggested by the allophony of L just covered — which they attribute to a rule, not to target-achievement-failure.
- Keating's VhV studies also take advantage of an uncrowded environment.

23. Dense targets

- These are found in
  - almost any segmental phonology
  - tonal phonology in a target-rich language like Mandarin Chinese.
- It is unlikely that they are achieved.
- Reasons not to achieve:
  - difficult "slaloms" of targets (Chinese, per Flemming and Cho)

  - lenition (why bother to achieve?). E.g. *get Bill*, where a glottal closure makes alveolar closure not very necessary for coda /t/ = [ˀt], [ʔᵗ].
- In the last class we will look at Harmonic Grammar as an approach to unachieved targets.

ARITHMETIC IS THE LANGUAGE OF PHONETICS

24. What arithmetic?

- Multiplication
  - Cleverly designed experiments yield not-so-pudgy snakes of data points, indicating a clear multiplicative relationship.
- Addition
  - This arises when the pudgy snake, extrapolated, does not pass through the origin.

25. One experiment that created pudgy snakes

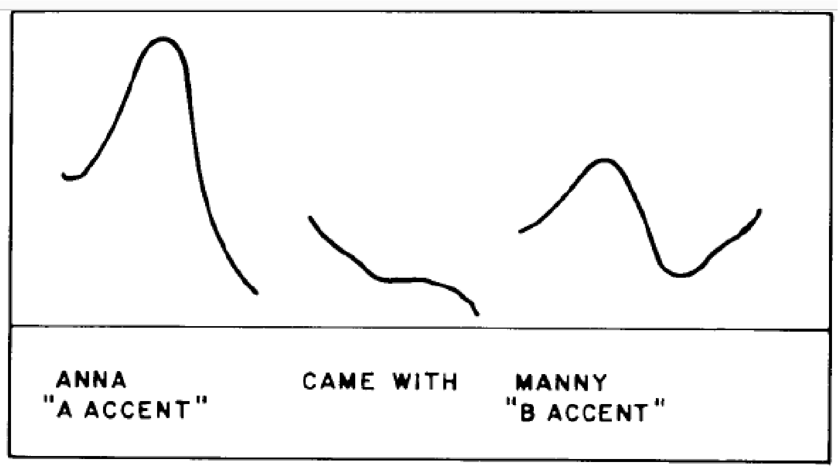


Figure 9
An F0 contour for *Anna came with Manny*, produced as a response to *What about Manny? Who came with him?*

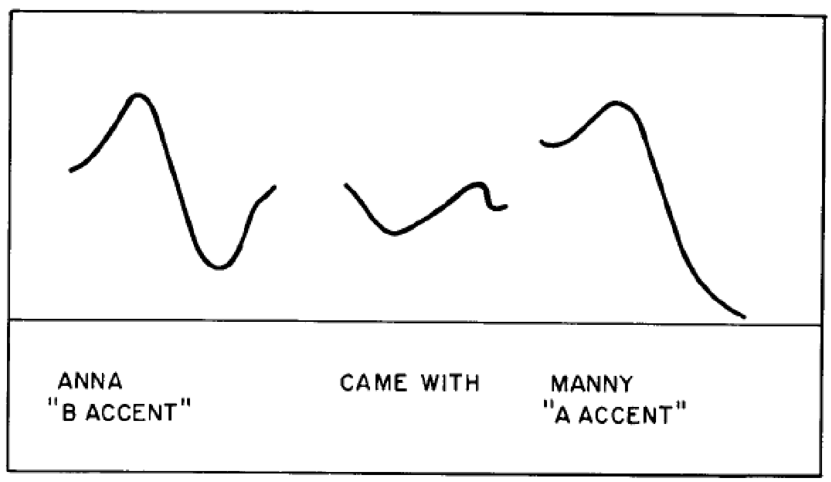


Figure 10
An F0 contour for *Anna came with Manny*, produced as a response to *What about Anna? Who did she come with?*

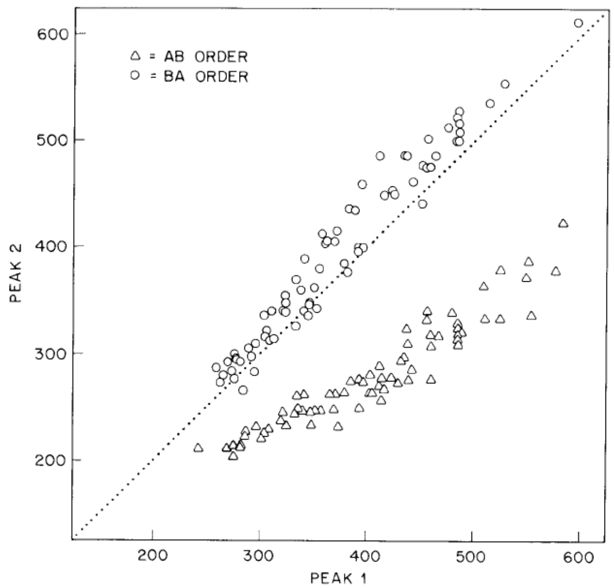
- Socrates: I can tell you roughly what these are tonally; fully-trained ToBI'ists please specify in full.

26. Experimental method

- Find incredibly-cooperative-and-sympathetic experimental subjects.
  - ➢ The authors themselves and colleagues at Bell Labs
  - ➢ I'm actually pretty sympathetic on this issue — rendering intonations out of context is very difficult.
- Give them little prompt-cards with emphasis-number from 1-10.

- This is key: *vary along a continuum*, to generate your snakes, hence detect lawful patterns — large range will lengthen the snake and de-sheep-ify it.

27. Sample pudgy snakes from this experiment (subject JP)



28. More snakes in the berry experiment

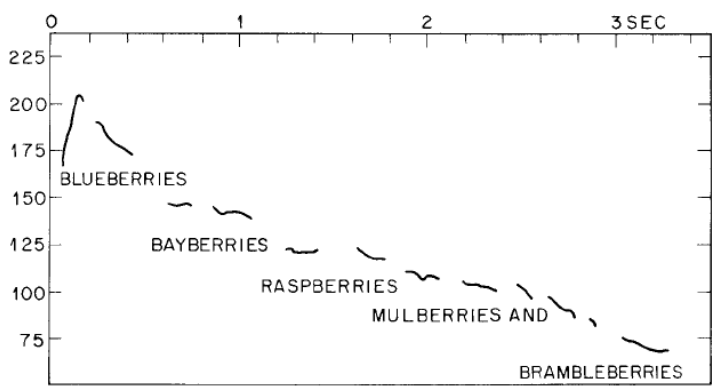


Figure 13
An F0 contour for the berry list *Blueberries, bayberries, raspberries, mulberries, and brambleberries*, produced with a sequence of step accents. Each step is smaller than the one before, so that the step levels appear to trace out an exponential decay.

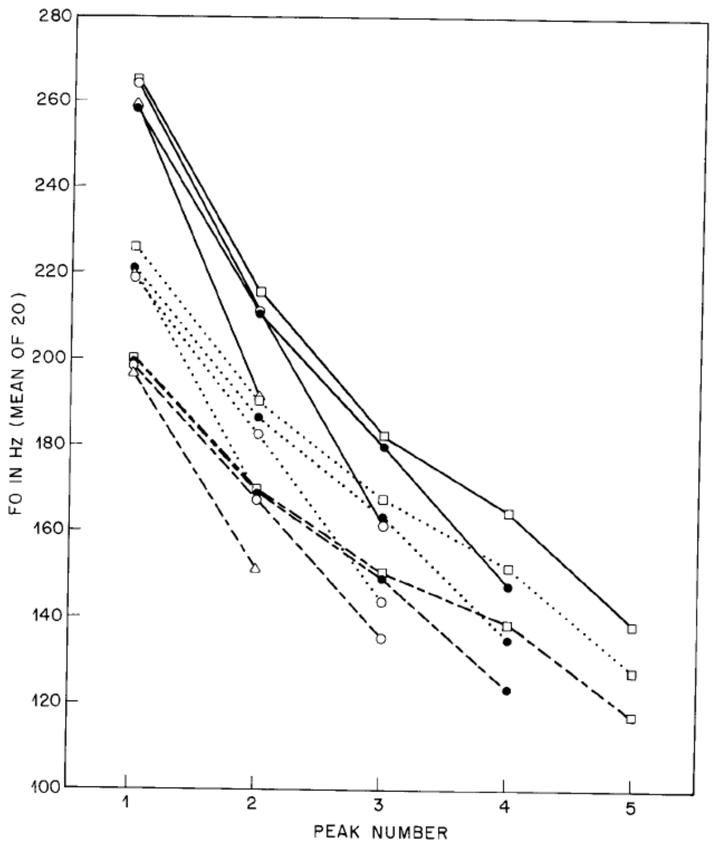


Figure 19
Downstep data (3 pitch ranges, 4 lengths) for subject DWS

29. The informativeness of this experiment

- They obtain an *exponential* relationship of successive pitch accents.
- This relationship is defined not on raw Hertz but on hertz-above-a-floor value.
- There is also a deviation at the end, justifying a phonetic rule of Final Lowering.

30. Regression as the key to developing models

- Put the hypothesized principles, relating {phonological entities, emphasis} to measurements, into arithmetic form.
- Do regression to obtain parameters with best fit.
- Seek areas of systematic error, revise and repeat, possibly doing further experiments.

31. The equations developed in Liberman and Pierrehumbert (1984)

- The model is a very simple one in which tone targets are computed from preceding tones.
- It uses fictional units of "high-ness", measured as Hertz above a reference line of height *r*

- It is fun to imagine a unit of measure, which Pierrehumbert (1980) facetiously calls the Amana.[4]

32. Defining the Amana

$T(P) = P - r$ defining the Amana by subtraction; P is physical pitch

33. The Downstep Rule

$$T(P_i) = s \cdot T(P_{i+1})$$

- Iterated downstep is a constant fraction, when measured in Amanas.

34. Answer-Background Scaling (for *Anna came with Manny*)

$T(P_A) = k * T(P_A)$

- It's the *same* ratio *k* despite difference in order, once the other rules are taken into account.

35. Final Lowering rule

$$P \rightarrow r + l \cdot (P - r) / \_\_\_\_ \$$$

- In some window to be defined near the right edge, a pitch target shrinks down by a factor of *l* when measured in Amanas.

36. Utter (confessed) ad hockery to give us a value of *r* from empirical data

$$r = f \cdot (P_0 - b)^e + d + b$$

where P_0 is the target in Hz of the first pitch accent, and d, e, f, and b are constants

- *b* is as low as you are willing to go
- *d* is positive
- they needed a curved function to fit the data and so added an exponent.
- They struggle with this, it emerges as the best fit to their data but is less obviously principle.
- Something like this is needed; it is the seat-of-the-pants theory of emphasis.

[4] An American brand name for household appliances, consisting entirely of sonorants and useful for intonational work. Amana

37. Are the constants of the system true constants of English intonational phonology?

- Measuring the universal constants is an important activity of physicists.
- We might hope that vaguely similar values obtain for different subjects.

Parameter values for fit of model 1

	s	k	l	f	e	b	d
MYL	.59	1.66	.68	.0059	1.67	64.3	4.9
DWS	.68	1.33	.77	.0049	1.63	81.3	9.7
JBP	.62	1.63	.68	.0049	1.64	111.9	21.8
KXG		1.59	.59	.0049	1.33	90.3	18.9

where

- *s* is the downstep constant
- *k* is the Answer-Background ratio
- *l* is the Final Lowering constant

- Conceivably there is dialect variation.
- Conceivably the values vary expressively (they say this later for *l*).
- Conceivably there was a Standard Bell Labs pronunciation system worked out unconsciously in the course mutual imitation.

REGISTRAL HIERARCHY AND SHAPE IN F0

38. The concept of autosegmental register

- This can be found in Goldsmith's (1976) dissertation *Autosegmental Phonology*
- In African intonation systems, the downstepping seen above in English also occurs; usually in alternating H and L: a H after L is not as high as the high that preceded it.
- Hence a *memory device*: "how high was H, the last time I uttered it?"
- The usual phonetic downstep rule is: "Lower the registers at the H-L transition."
- Exercise: deriving the various ways to say /ówà ówà/ in Etsako.

39. Downstep in Japanese

- Seeking precision (since it's not triggered by any H L transition), Pierrehumbert and Beckman call it *catathesis*.
- It is triggered by any pitch accent (H*+L)
- See figure above for the effect of catathesis on the final L boundary tone.
- It is detectible by the minimal pair method: measuring peaks on the same items as they occur after accented or unaccented words.
  - This chart also varies focus.

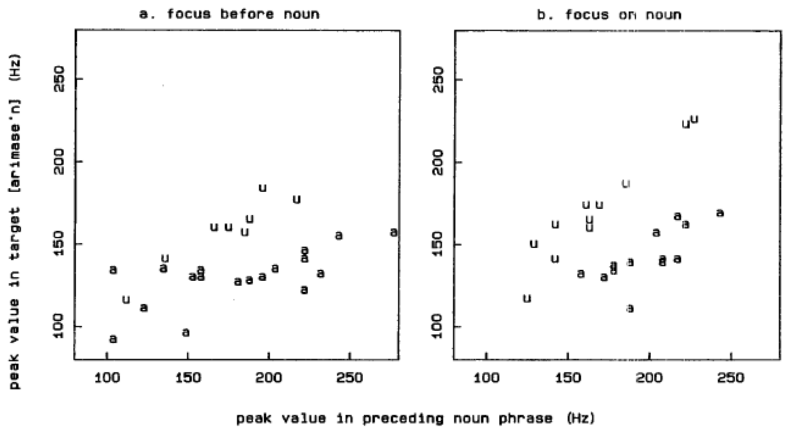


Figure 3.1
Value of the f_0 peak in *arimase'n* plotted against the peak in the preceding noun in utterances from the "amaimame" set produced by speaker TS. Plotting characters: "a" = accented noun *mame'*, "u" = unaccented noun *ame*. Panel (a) shows tokens with focus on the adjective preceding the accented or unaccented noun and (b) tokens with focus on the noun itself.

40. The key finding for scaling in Japanese

- *Every* tone is downstepped after an accent, each in suitable proportion.
- How to get this result? Basically, *downstep must change the Amana, not the tones*.
  - ➢ The metaphor of stretchable graph paper, scaled 0 - 1, on which each tone is plotted with its own inherent number.
  - ➢ e.g., if, one such graph paper, accented H* is 1 Amana, unaccented H is .9 Amanas, and L is .3 Amanas, the basic shape of pitch patterns in accentual phrases will remain the same.
- Hence, it appears that phonetics has hidden structure, like the other components of the grammar.