ETAP4@UMass

Modeling ‘Elbows’ in F0 Contours: Phrase Accents in English
Edward Flemming (flemming@mit.edu)
Massachusetts Institute of Technology

The problem with Elbows: an example
- Pierrehumbert (1980) posits leftward spreading of L- in H*L-H% and H*L- L% tunes to explain why F0 does not interpolate from H* to the end of the phrase.

PvH (Hz) H* L- H%

- Two hypotheses concerning the timing of the onset of L- (Pierrehumbert 1980):
  1. L- occurs at a fixed interval after H*
  2. L- is aligned to the end of the nuclear-accented word.
- To test these hypotheses we have to locate L-:
  - The correlate of L- is an ‘elbow’ or inflection in the F0 trajectory
  - ‘it was very easy to decide where the L- was located.’ (Pierrehumbert 1980:86)

Testing hypotheses about the timing of L-:
- Predictions of the two hypotheses: What should happen to the time constant T when the duration of the interval between H* and the end of the word varies?

T<0.034
L- aligns to end of word
T<0.044
L- occurs at a fixed interval after H*

- T is predicted to be a linear function of H*-to-end duration, with intercept = 0
- Slope of the line is 1/n, where n is the number of time constants to reach the L-
target.

Identifying elbows through analysis-by-synthesis
- Analysis-by-synthesis of F0 trajectories:
  - Rather than identifying elbows using general-purpose algorithms, then modeling the results (e.g. del Giudice et al 2007, Reichel & Salvete 2015),
  - ‘Elbow’ targets should be inferred in the process of modeling F0 trajectories.
- Model of H*L-(T%) production:
  - The transition from H* to the first L- target is realized as the response of a critically-damped linear second order system (‘spring-mass system’) to a step input.
  - The transition from the first L- target to the second is the response of the same system to an input linear transition between the two targets,

Results I – production model
- The critically-damped model does not fit all speakers/utterances well
  - Problem: damped ‘spring-mass’ models have peak acceleration at movement onset, but this is not true of all H*L- transitions

Results II – timing of L-
- T tends to increase as duration from H* to end of word increases (β = 0.11, t = 6.5)
- So the interval between H* and L- is not fixed, but L- does not track word end either –β = 0.11 would imply that target is achieved at 9T, also intercept > 0 (β = 0.012, t = 3.4)
- This pattern could represent a compromise between a preferred value for T and a preference to keep L- within the accented word, but there is a lot of variability.

The Data
- Recordings from Barnes, Veilleux, Bragos, & Shattuck-Hufnagel (2010)
- 25 two-word phrases in a context designed to elicit H* L- H% melody, with H* on the first word.
- First word: vary the number and length of syllables following primary stress
  1. a lien, lànnol, Jillian, Marilyn, minimum
  2. binmatory, pambilony, cernynery, culinary, pulmonany
  3. criminally, sèrialy, ternimaly, minimally, nominally

P (t) = \begin{cases} 0 & \text{if } t \leq 0 \\ \frac{1}{2} \left( 1 - e^{-\frac{t}{T}} \right) & \text{if } t > 0 \end{cases}

- (1) Experimental: George is a thoughtful sort of divorce lawyer—I go to him whenever I need a pambilony ruminator.
- Subject: A pambilony ruminator?????? (H* L-H%) I thought he was figuring out your plumbing problem!
- 15 speakers (11 female), each produced 4 repetitions of the materials.
- 239 utterances excluded due to errors, discontinuities, pitch tracking problems.
- Tracked F0 with Praat (Boersma & Weenink 2018), segmented the pitch contour from F0 peak (H*) to onset of the final rise, and fitted the tone realization model using non-linear least squares (nL.5B (R Core Team 2016)).


Download this poster: web.mit.edu/~flemming/www/paper/elbows.pdf

Thanks to Jon Barnes, Alejna Bragos, Stefanie Shattuck-Hufnagel and Nanette Veilleux for sharing their recordings, and to Run Chen for research assistance.