

# Looking beneath the surface: Why AM theory needs to be revised

Laura C. Dilley, Ohio State University, [dilley@ling.ohio-state.edu](mailto:dilley@ling.ohio-state.edu)

## Overview

- Evaluating how well a theory explains facts about perception and production requires a clear understanding of the theory's assumptions about the phonology-phonetics interface.
- Question:** For the autosegmental-metrical (AM) theory of tone [1-4], how are physical F0 patterns assumed to be related to phonological representations? Moreover, what are the implications of this view for the adequacy of AM theory with respect to tonal perception and production?
- It is demonstrated that seldom-recognized assumptions about the phonological vs. phonetic status of *relative tone height* lead to serious and significant problems with basic description under the theory. A revised version of the AM framework is proposed which eliminates these problems, while more effectively dealing with facts about tonal perception and production.

## Phonetic models of intonation

- Ladd [5], p. 37 states: "Unconstrained gradient variability of prominence and other pitch range parameters is, in my opinion, the most serious empirical weakness of a great many quantitatively explicit models of F0."
- There has been almost no work examining the theoretical implications of *unconstrained gradient variability* (UGV) in phonetic models of AM theory. The implications of UGV are explored for two models: Pierrehumbert (1980) [3], or *P80*, and Pierrehumbert and Beckman (1988) [4], or *PB88*.

### The model of P80

- Approximately 10 phonetic rules are proposed to adjust the heights of tones ([3], pp. 144-159). Adjustments in tone height are made with respect to (1) the speaker's pitch range and (2) other tones. Some phonetic parameters in the model:

$$f_0(T) = F_0 \text{ of a tone } T$$

$$p(T) = \text{"prominence"} \text{ of a tone } T$$

$$n, k = \text{unspecified parameters } (0 < n < k < 1)$$

- Example.** Consider how the F0 levels of tones in a H\* L+H\* sequence are determined. The F0 values are determined according to Eqns. I and II (rewriting H\* L+H\* as H<sub>1</sub> L+H<sub>2</sub>):

$$(I) \frac{f_0(L+)}{f_0(H_1)} = n \cdot \frac{p(H_1)}{p(L+)}$$

$$(II) \frac{f_0(L+)}{f_0(H_2)} = \frac{n}{k} \cdot \frac{p^2(H_1)}{p(L+)p(H_2)}$$

- Problem:** UGV in values of *p(T)* causes H\* L+H\* to generate either of the F0 contours in Figure 1, as well as others.\* Similar problems obtain for most tonal sequences in the model.

\*The contour in Fig. 1b is generated when the right-hand sides of eqns. (I) and (II) are greater than 1/n and k/n, respectively (Dilley [6], pp. 153-154, FN 49).

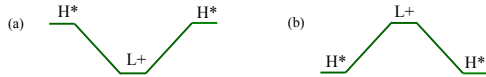


Figure 1.

### The model of PB88

- Multiple phonetic parameters are proposed to affect the heights of tones, all of which relate to the speaker's pitch range. Some of these parameters are as follows:

$$p(T) = \text{"prominence"} \text{ of a tone } T; \text{ ranges from } 0 \text{ to greater than } 1 \text{ (unbounded)}$$

$$h = F_0 \text{ of high tone line}$$

$$l = F_0 \text{ of low tone line}$$

- Example.** Consider how the F0 levels of a H\* L+H\* sequence is determined. The F0 values are determined by *p(T)*, *h* and *l*, as in Figure 2.

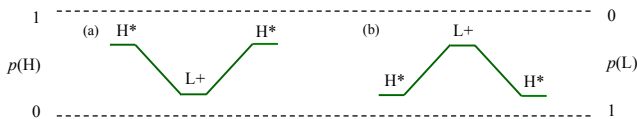


Figure 2.

- Problem:** UGV in values of *p(T)* causes H\* L+H\* to correspond to either of the contours in Figure 2, as well as others.\* Similar problems obtain for all tonal sequences under the model.

\*The contour in Fig. 2b is generated when 1-p(H) > p(L) for adjacent L and H in H<sub>1</sub> L+H<sub>2</sub> (Dilley [6], pp. 66).

## Summary

- These results show that UGV in phonetic models complementing AM theory leads to the following problems:
  - Too many F0 contours arise from a given tonal sequence (*overgeneration*).
  - A given F0 contour could correspond to many underlying representations (*indeterminacy*).
- Next, we consider *why* these negative consequences arise in AM theory, what their phonological implications are, and what can be done about them.

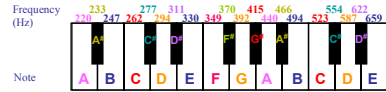
## Phonological implications

- The descriptive problems in phonetic models complementing "standard" AM theory are due to the treatment of *relative tone height* in the phonetics-phonology interface.
  - e.g., Given a HL sequence, H is not restricted to be higher than L.
- Question:** Why is relative tone height not sufficiently restricted? The answer can be traced to *phonological* assumptions. Recall two tenets of autosegmental theory [1]:
  - Tones and segments are represented on different tiers.
  - Tones are "bundles of features," just like segments. cf. H = [+high, -low], L = [-high, +low]
- Answer:** Problems arise because claim (2) entailed the assumption that tones are not inherently relational. *There is no formal phonological sense in which H is "higher than" L.*
- Question:** In order to eliminate UGV, should the *phonetic* component of AM theory be modified, or the *phonological* component?
  - Answer:** The *phonological* component should be modified, for at least two reasons:
    - Perception.** Merely modifying the phonetic component does not account for the fact that relative tone height is used in perception to recover tonal representations (e.g., [7]).
    - Cross-linguistic evidence.** Relative tone height is now known to be contrastive in many languages, and accounts of this based on HLH sequences triggering phonetic lowering are descriptively inadequate [8].

## Drawing phonological inspiration from musical representation

- In deriving a revised AM approach, it is worthwhile to consider what is known about the representation of tonal categories in other domains, such as music. Three facts are relevant:

- Representational categories for tone in music are inherently relational [9].
  - Musical tone categories are based on ratios of frequencies, e.g. 1 semitone =  $^{12}\sqrt{2} = 1.0594$ .



- Musical melodies are recognized by the sequence of ratio relationships among the notes. For example, Figures 3(a) and (b) correspond to the tune "Happy Birthday" in the keys of C and F, respectively.

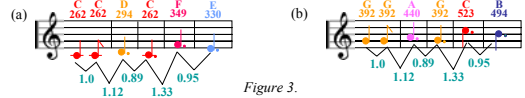


Figure 3.

- The mapping from physical frequencies to representations is determined by culture and experience [9]. Moreover, the perception of physical frequencies in terms of musical representations is categorical in nature (e.g., [10]).
- There are "layers of representation" for musical melody varying in complexity (e.g., [11]).
  - Up-down pattern* = whether successive notes are higher than or lower than other notes
  - Interval* = distance between notes, in terms of a specific frequency ratio or range of ratios
  - Scale* = relation between a note and a tonic referent note in a particular key

## A revised AM theoretical framework: Proposals

- Tones and segments are represented on different autosegmental tiers [1].
- Tones associate with respect to timing positions specified in metrical grids [2].
- Tonal features are defined as frequency ratios or ranges of ratios (called *tone intervals*, *I*). Two kinds of relationships may be defined:
  - Syntagmatic:**  $I_{1,2} = T_2/T_1$ 
    - If  $I_{1,2} > 1$  (*higher*), then  $T_2 > T_1$  (Fig. 4a)
    - If  $I_{1,2} < 1$  (*lower*), then  $T_2 < T_1$  (Fig. 4b)
    - If  $I_{1,2} = 1$  (*same*), then  $T_2 = T_1$  (Fig. 4c)
  - Paradigmatic:**  $I_{r,0} = T_0/T_r$ , where  $T_r$  is the *tonic referent tone*.
- Further language-specific restrictions on values of *I* account for differences in the size of a rise or fall, and the position of a tone in the speaker's pitch range.
- The tonal system of English and other intonation languages is proposed to be comparatively simple, defining only syntagmatic tone intervals. In contrast, the tonal systems of many languages of Africa and Southeast Asia are proposed to be more complex, defining both syntagmatic and paradigmatic tone intervals.
- This approach eliminates problems of UGV from AM theory by restricting the relative heights of tones, thereby improving descriptive and explanatory adequacy for tonal facts.

## Conclusions

- A little-known problem with standard AM theory concerns the fact that unconstrained gradient variability in phonetic parameters leads to a breakdown of descriptive power.
- These problems ultimately stem from a strong phonological assumption in the original proposals of autosegmental theory which has since been discounted.
- A revised version of AM theory is proposed in which relative tone height is treated as phonological. This eliminates problems with unconstrained gradient variability, thereby restoring both descriptive and explanatory power.

## References

[1] Goldsmith, J. (1976). *Autosegmental phonology*. Ph.D. dissertation, MIT, Cambridge, MA.  
 [2] Liberman, M. & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, 8(2), 249-336.  
 [3] Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation*. Ph.D. dissertation, MIT, Cambridge, MA.  
 [4] Pierrehumbert, J. & Beckman, M. (1988). *Japanese Tone Structure*. Cambridge, MA: MIT Press.  
 [5] Ladd, D. R. (1990). Metrical representation of pitch register. In J. Kingston & M. Beckman (Eds.), *Papers in Laboratory Phonology 1: Between the Grammar and Physics of Speech* (pp. 35-57). Cambridge: Cambridge University Press.  
 [6] Dilley, L. (2005). *The phonetics and phonology of tonal systems*. Ph.D. dissertation, MIT, Cambridge, MA.  
 [7] Wong, P. C. M. & Diehl, R. L. (2001). Perceptual normalization for inter- and intra-speaker variation in Cantonese level tones. *Journal of Speech, Language, and Hearing Research*, 46, 413-421.  
 [8] Soltes, K. (1999). *The Geometry and Features of Tone*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington Publications in Linguistics 13.  
 [9] Burns, E. M. (1999). *Intervals, scales, and tuning*. In D. Deutsch (Ed.), *The Psychology of Music* (2nd ed., pp. 215-264). San Diego: Academic Press.  
 [10] Burns, E. M. & Ward, W. (1978). Categorical perception - phenomenon of auditory evoked: Evidence from experiments in the perception of melodic musical intervals. *J. Acoust. Soc. Am.* 63(2), 456-468.  
 [11] Handel, S. (1989). *Listening: An Introduction to the Perception of Auditory Events*. Cambridge, MA: MIT Press.