Perceptual Robustness of the Tonal Center of Gravity for Contour Classification

Nanette Veilleux, Jonathan Barnes, Stefanie Shattuck-Hufnagel, Alejna Brugos
Simmons College, Boston University and Massachusetts Institute of Technology
veilleux@simmons.edu, jabarnes@bu.edu, stef@speech.mit.edu, abrugos@bu.edu

In traditional phonemic analysis, phonologists rely on speaker intuitions regarding meaning differences to demonstrate phonological contrast in sound patterns. The discreteness of these differences, typically taken for granted, serves as the basis for inferences regarding the discreteness of phonological distinctions. In intonational phonology, however, analogous native-speaker intuitions are often substantially less sharp. Intonational meaning is notoriously slippery, rarely truth-conditional, and frequently confounded by informants with aspects of affect and paralanguage. Native speakers and linguists alike often stumble over seemingly innocent questions, such as whether two contours (e.g., L+H* L-H% and L*+H L-H%) "mean the same thing". Because of this, the traditional King's Highway to phonemic contrast, the minimal pair test, is often of little use in the study of intonation, leading researchers [e.g., 5, 9] to seek demonstrations of phonological contrast by other means, such as the identification of sharp boundaries between intonational categories in perception or production. Pierrehumbert [8: 59-63] explicitly advocates discreteness in realization as evidence for phonological contrast, and from there infers distinctions in meaning, rather than the other way around.

Locating categories in perception or production naturally requires a variable expressing the relevant boundaries. In AM intonation models, this usually involves scaling or alignment of F0 turning points (maxima, minima, "elbows", hereafter TPs) [3,8]. While few would posit exceptionless correspondence between phonological tones and TPs [6: 103-107], most arguments for contrast in intonation today presuppose some level of equivalence. This assumption yields tremendous successes as a laboratory tool [1, *et seq.*], but nonetheless encounters difficulties as a model of "what matters" in performance. If TP alignment provides a primary cue to tonal representation in phonology, then it should be robustly recoverable in both perception and acoustic analysis. An increasing body of evidence, however, suggests that this is not case; instead, apparently crucial TPs are either difficult to locate, as in ambiguous 'elbows' in the F0 contour [2], or absent, as in voiceless regions. Aspects of global contour shape beyond TP alignment are also known to play a role in categorization [4, 7]. This paper presents a new model of tonal alignment that distinguishes contrasting categories in English more robustly than TP-based models, while capturing the influence of contour shape on perception as well. At the heart of this model is the notion of Tonal Center of Gravity (TCoG), a distillation of contour shape returning the location where the "central bulk" of a pitch excursion lies. TCoG is calculated as an average of timepoints in an interval weighted by their F0 values. (See figure 1.) Points with higher f0 are thus more influential then those with lower f0, "pulling" TCoG in their direction.

To investigate the robustness of TCoG as a marker of tonal alignment contrasts, we elicited tokens of L+H* L-H% and L*+H L-H%, contours from 6 native English speakers. Productions were labeled for TPs and TCoG, and automatically categorized by logistic regression analyses representing all models. To explore robustness, we then subjected both TP labels and interval edges for TCoG to random perturbation drawn from a Gaussian source, Classification was thereupon repeated. This procedure was repeated 30 times at 3 levels of added label "noise". For noiseless data, TCoG categorized better than either the "Low" or the "High" TPs alone, and approximately as well as a model combining these two. With perturbed labels, TCoG significantly outperformed all other models. TCoG is thus less vulnerable to ambiguities in the F0 contour, and we suggest also a more reliable perceptual cue to contrasts between contours similar in shape, but differing in meaning.

Figures and Tables

Figure 1: Tonal Center of Gravity (TCoG) is an f0 weighted average of time points, representing the center of the "bulk" or "mass" of f0 over an interval.

$$TCoG = \frac{\sum (time \times f0)}{\sum f0}$$

Table 1: Correct detection of L+H* L-H% vs.L*+H L-H% when the rise and peak are perturbed by Gaussian noise based on multiples of the standard deviation for each speaker's rise/peak location. Figures represent averages over 30 trials for each noisy condition.

Level of noise	0	1	2	5
Rise (L)	79	70	65	60
Peak (H)	83	72	66	57
Rise + Peak (LH)	90	78	70	61
TCoG	90	85	80	72

References

[1] Arvaniti, A., Ladd, D. R., & Mennen, I. (1998). Stability of tonal alignment: the case of Greek prenuclear accents. *Journal of Phonetics,* 26: 3-25.

[2] Barnes, J., Brugos, A., Shattuck-Hufnagel, S. & Veilleux, N. (2008). Turning Points, Tonal Targets, and the English L- Phrase Accent. Paper presented at Experimental and Theoretical Advances in Prosody, Cornell University.

[3] Bruce, G. (1977). *Swedish word accents in a sentence perspective.* Lund: Gleerup.

[4] D'Imperio, M. (2000). *The Role of Perception in Defining Tonal Targets and their Alignment*. Doctoral dissertation, The Ohio State University.

[5] Kohler, K. (1987). Categorical pitch perception. *Proceedings of International Congress of Phonetic Sciences 11* (Tallinn), vol. 5: 331–333.

[6] Ladd, D.R. (1996/2008). *Intonational Phonology*. New York: Cambridge University Press.

[7] Niebuhr, O. (2007). The Signalling of German Rising-Falling Intonation Categories – The Interplay of Synchronization, Shape, and Height. *Phonetica*, 64: 174-193.

[8] Pierrehumbert, J.B. (1980). *The Phonetics and Phonology of English Intonation*. Doctoral dissertation, MIT.

[9]Pierrehumbert, J. B. & Steele, S. A. (1989). Categories of tonal alignment in English. *Phonetica*, 46: 181-196.