

Interpreting Patterns of Variability in the Realization of English Intonation Contours

Jonathan Barnes, Alejna Brugos

Romance Studies & Applied Linguistics,
Boston University, Boston, MA, USA

Nanette Veilleux

Computer Science and Informatics, Simmons
College, Boston, MA, USA

Stefanie Shattuck-Hufnagel

Research Lab of Electronics, Massachusetts Institute of
Technology, Cambridge, MA, USA

Introduction: Toward a global model of intonation perception

From F0 contours to tone categories in perception:

Theories of mapping from raw F0 curves to phonological tone categories traditionally rely on certain key points identifiable within the contour, usually local maxima, minima, or elbows [4].

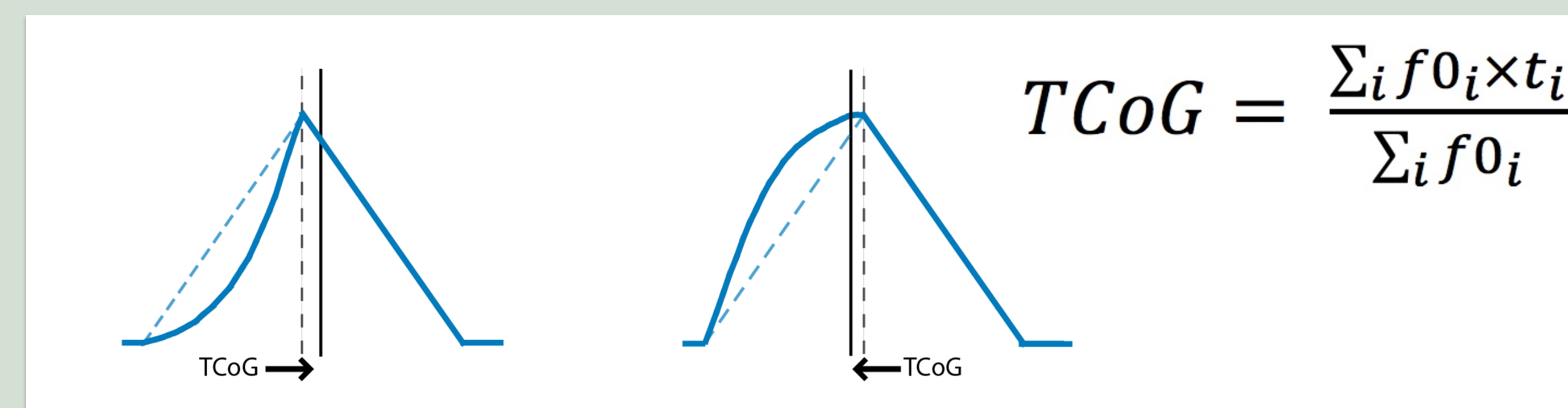
Advantage: achieves immediate data reduction by effectively discarding the rest of the contour for purposes of further processing.

Disadvantage: assumes those points are in fact the only things that matter to listeners in perception.

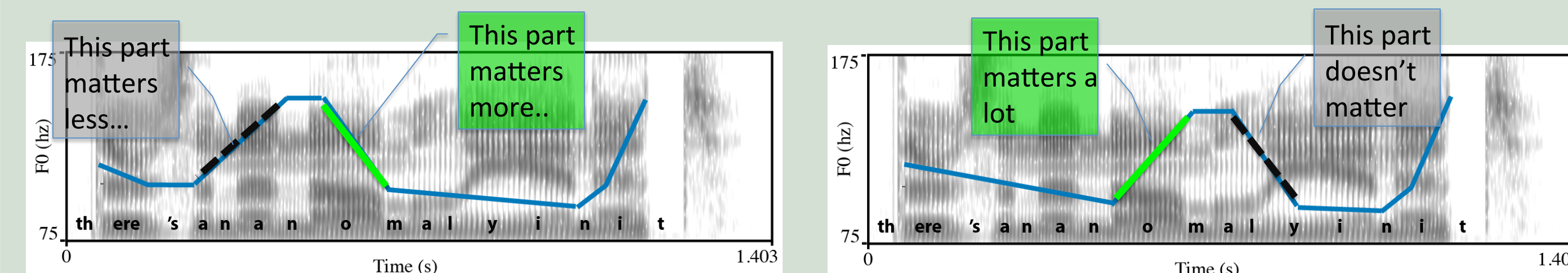
Two problems with using key points in F0 contour

- 1) **F0 contour shape matters:** Aspects of contour shape not easily reducible to the location of a single point or points on the contour are known to matter in perception [2,4,8,9]. E.g., Curvature of F0 rise shapes judgments of category membership (L+H* vs. L*+H) [2]

A global model of tone perception that integrates F0 information over time can account for these facts



- 2) **Aspects of segmental skeleton “hosting” F0 contours matter too:** The same raw F0 contour shape is processed differently as a function of how it aligns with key portions of the segmental skeleton. E.g., listeners are sensitive to different parts of a plateau-shaped contour depending on how it is aligned w/r/t the accented syllable [2]



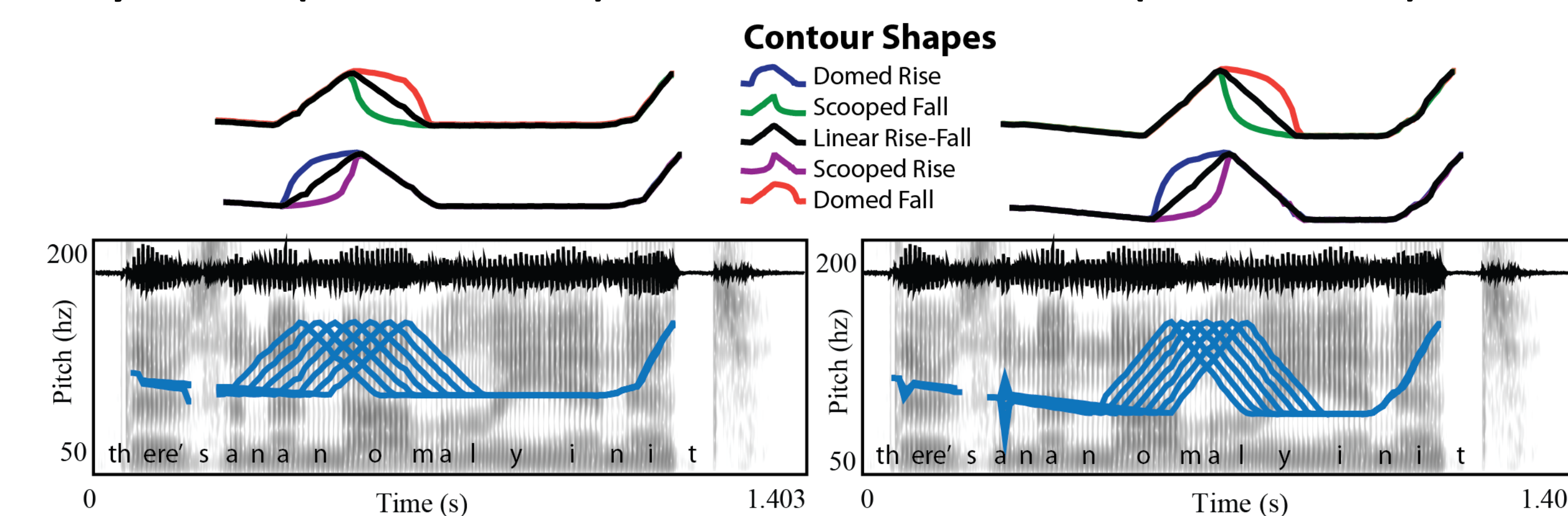
Goals of the present study:

To determine how F0 contour shape and segmental context influence perception of categories in a global model of intonation.

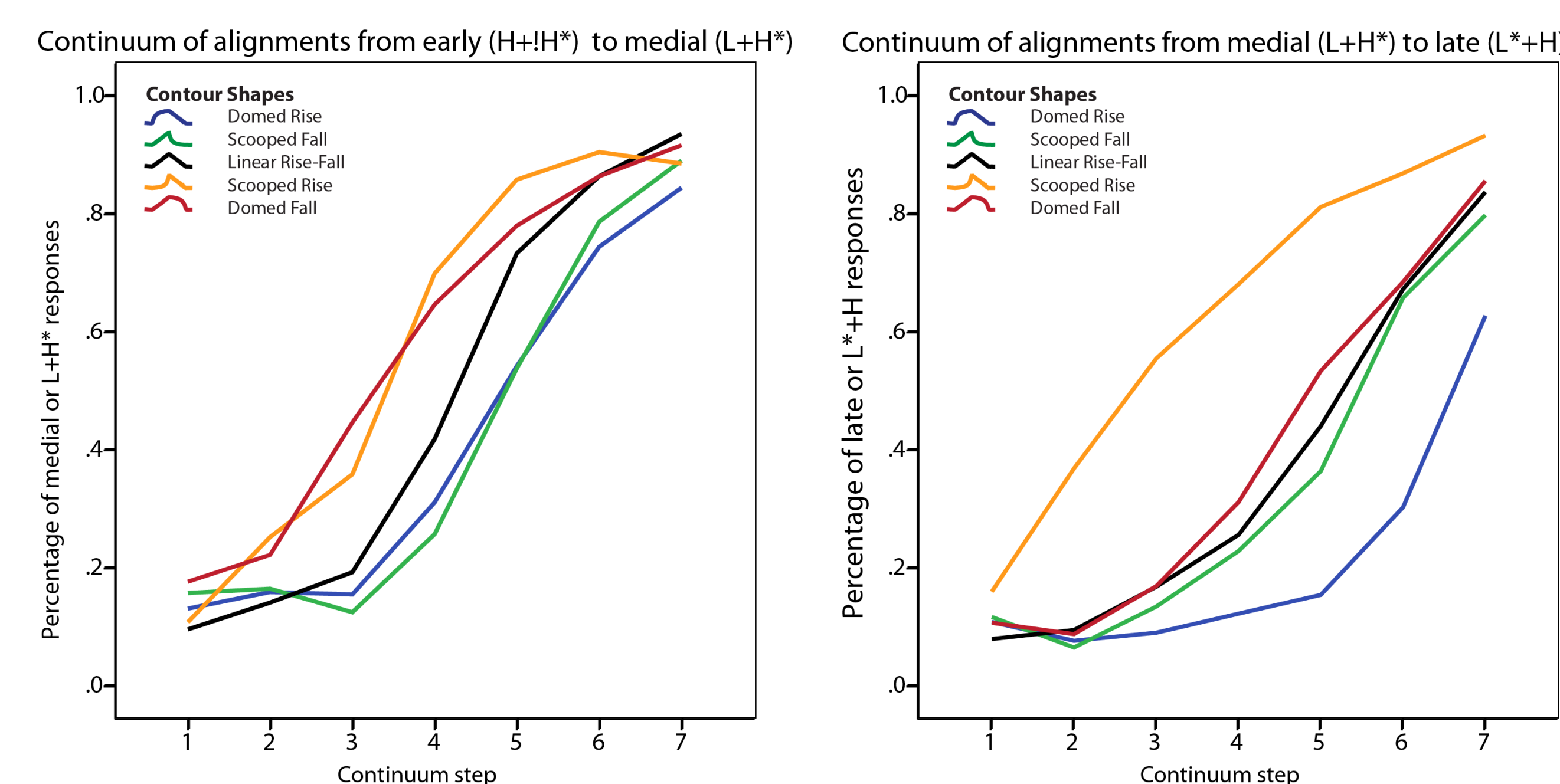
Experiment

- 27 subjects heard two pairs of contours in AXBX order
- A and B are endpoints of a timing continuum; X a contour from that continuum.
- 5 contour shapes X 7 alignments x 2 timing continua (run in 2 sessions)
- Stimulus utterance: *There's an anomaly in it*
- Perception task: Which pair “matches”?

Early-to-Mid (H+!H* to L+H*) continuum Mid-to-Late (L+H* to L*+H) continuum



Results



Early-to-mid continuum: scooped rises/domed falls bias listeners toward L+H*; domed rises/scooped falls do the opposite.

Mid-to-late continuum: scooped rise biases listeners toward L*+H; domed rise toward L+H*; scooped and domed falls similar to linear rise/fall.

Discussion

- Early-to-mid continuum: standard TCoG predicts results for both rises and falls. Listeners respond the same to different shapes if effect on TCoG is same
- Mid-to-late continuum: Standard TCoG also predicts results due to the rise
- However, changes to the fall in mid-to-late continuum have no effect on results

How to account for this?

- Alignment of F0 contour with segmental skeleton affects perceptual “weighting” of individual F0 samples.
 - Portions of the contour realized over segmental “sweet spots” matter a lot in perception.
 - Portions of the contour realized over other segmental positions matter less.

For more information on TCoG, please visit:
<http://blogs.bu.edu/prosodylab/publications/>

Discussion (continued)

Question still to be resolved: What exactly is a segmental sweet spot?

Candidate “sweet spots”:

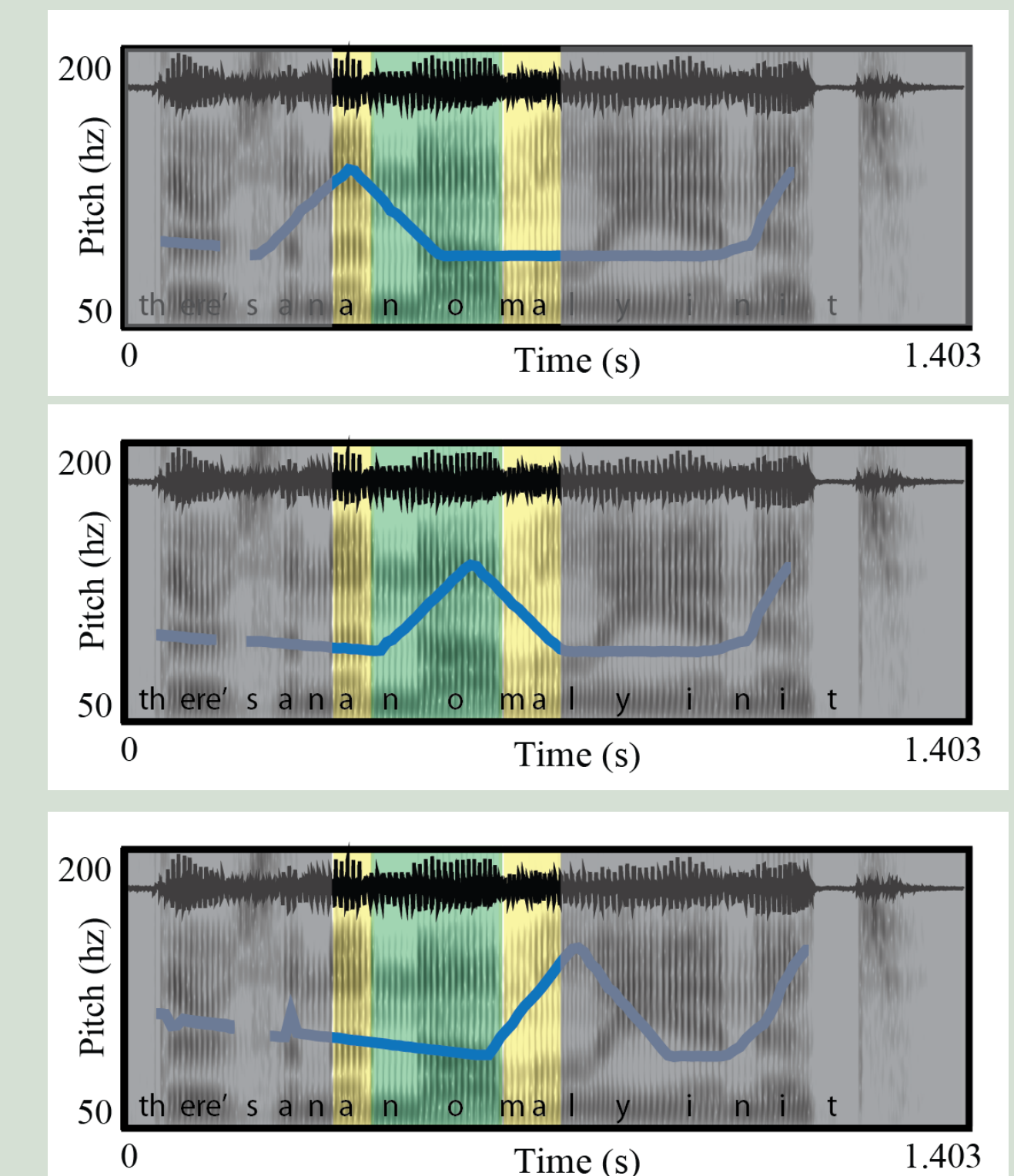
- accented syllables, nuclei
- higher sonority segments

Candidate “dead zones”:

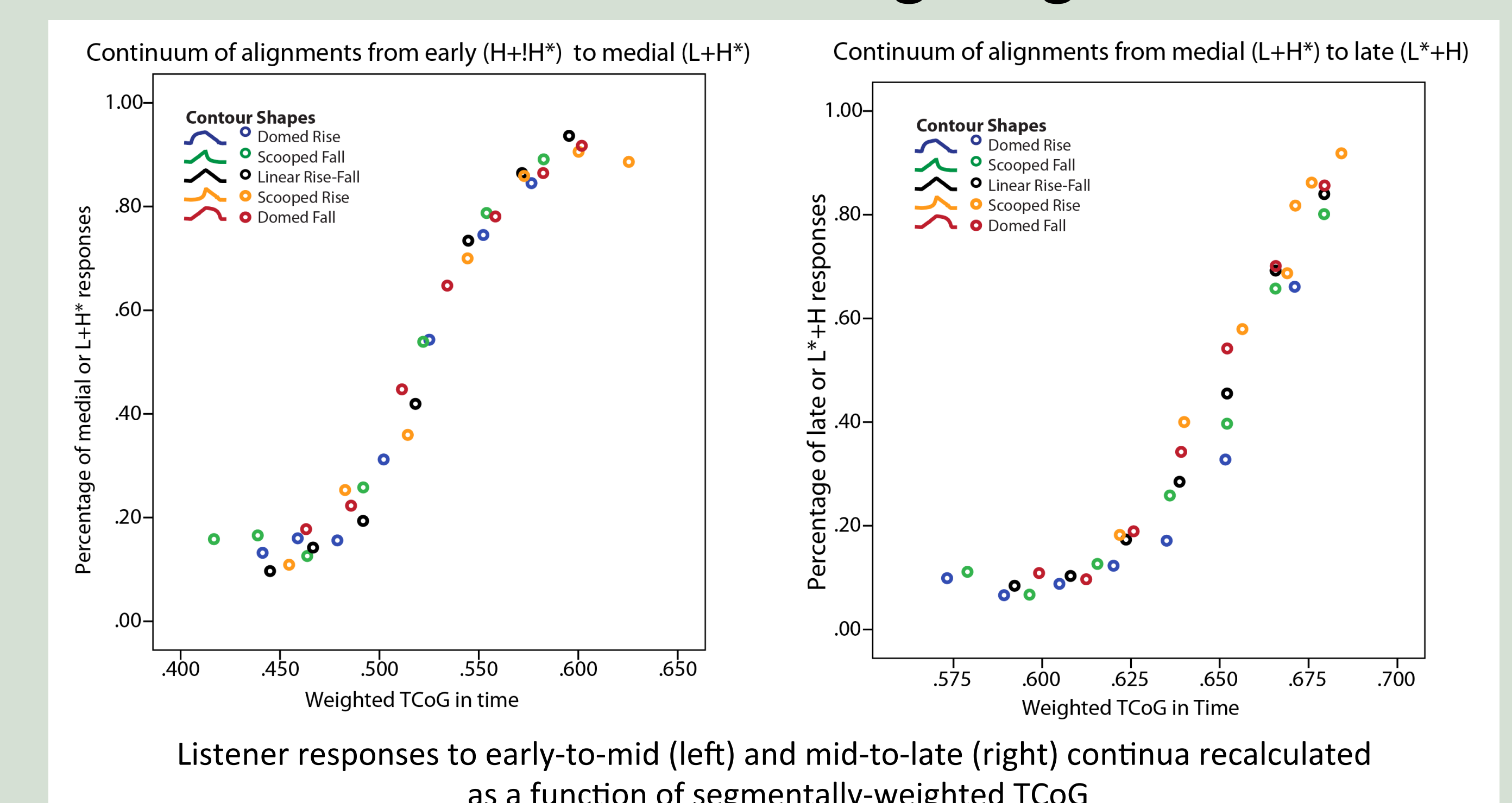
- unaccented syllables
- lower sonority segments

Sample “weighting scheme” implemented in TCoG

- Heaviest weight given to F0 samples in pitch-accented V and sonorant onset
- Declining weight to samples farther from center of this window



Recalculated results with weighting



Conclusions

- Perceived tonal categories change with alignment of TCoG: Segmental sample weighting accounts for differences in listener response to contour shape between the two alignment continua.
- Discussion of F0 “targets” minus global contour shape is meaningless.
- Discussion of contour shape minus segmental backdrop also meaningless.
- In perception, at least, discussion of pitch contours without reference to the segments that host them is deeply inadequate.

Acknowledgements

Supported by NSF grant numbers 0842912, 0842782 and 0843181.

References

- [1] Barnes et al. (2012). In: O. Niebuhr (ed.), *Prosodies –Context, Function, Communication*, 93-118. Berlin/New York: de Gruyter.
- [2] Barnes et al (2012). *Laboratory Phonology* 3(2), 337-383.
- [3] Barnes et al (2014) *Speech Prosody* 7, 1125-1129.
- [4] Bruce, G. (1977) *Travaux de l'Institut de Linguistique de Lund* 12.
- [5] Lund, Sweden: CWK Gleerup.
- [6] D'Imperio, M. (2000) Thesis, Ohio State University.
- [7] Flemming, E. (2008) In M. Embarki and C. Dodane (eds.), *La Coarticulation: Indices, Direction et Representation*.
- [8] Gordon (2001) *Studies in Language* 25: 405–444.
- [9] Knight, R. (2008) *Language & Speech* 51, 3: 223-244.
- [10] Niebuhr, O. (2007) *Phonetica* 64(2) 174-193.