

Prosodic Encoding of Speaker Attitudes

Johannes Heim

University of Greifswald

Introduction: In this talk, I present experimental evidence for correlations between the shape of the sentence-final rise in English and the perceived speaker confidence as well as their response expectation. Specifically, I report the findings of a complex perception study where native speakers of Canadian English rated speaker confidence and response expectation for rising declaratives. These declaratives varied in excursion and duration of the final pitch movement. The reported interpretability of the shape of a sentence-final rise suggests that previous characterizations of intonational meaning that only rely on bitonal distinctions (e.g. Pierrehumbert & Hirschberg 1990; Bartels 1997; Truckenbrodt 2012) need to integrate phonetic detail to model the encoding of speaker attitudes. The reported findings also point to the possibility of an independent perception of excursion and duration, which has been the subject of controversy in previous investigations of prosody (Rietveld & Gussenhoven 1978; Ladd & Mourton 1997; Armstrong *et al.* 2015).

Problem and proposed solution: A sentence-final rise encodes functions such as surprise, continuation, uncertainty, and contrast (O'Connor & Arnold 1973; Pierrehumbert & Hirschberg 1990). This multi-functionality is difficult to incorporate under a uniform analysis. Consequently, the burden of interpretation is shifted to context for the disambiguation of the different functions of a rise (Arvaniti *in press*). Contrary to this consensus in the literature on intonational meaning, I propose that the speaker uses intonation to signal their attitudes and intentions, which are readily interpreted by the addressee. The different functions associated with intonational meaning can be decomposed into signalling the speaker's commitment and their request for an engagement with their utterance by the hearer. COMMITMENT, which is encoded by the duration of the nuclear tune, expresses the extent to which the speaker is willing to publicly commit to the truth of a proposition. ENGAGEMENT, which is encoded by the excursion of the nuclear tune, expresses the extent to which the speaker signals that the hearer needs to engage with the utterance. Previous studies that have investigated the shape of the final contour have found that prosody is relevant for encoding questions (Fletcher & Loakes 2010), floor holding (Ritchart & Arvaniti 2014), speaker confidence (Pon-Barry 2008) or expertise (Tomlinson & Fox Tree 2010). The majority of the experimental literature has focused either on the amplitudinal or the temporal aspects of the sentence-final contour. Rietveld & Gussenhoven (1978) propose that these dimensions are dependent, Armstrong *et al.* (2015) propose that both aspects are perceived independently by the hearer.

Methods and Results: Empirical support for the hypothesis that the shape of a sentence-final rise encodes speaker attitudes comes from the findings of a two-part perception study which tested the effects of variation in pitch excursion and duration for ratings of speaker confidence and response expectation. Forty native speakers of Canadian English rated eighteen audible rising declaratives manipulated by two degrees of excursion and three degrees of excursion on a 5-point Likert scale. These 108 items were presented together with 108 control items with a falling contour and 108 filler items with monotonous or whispered contours. In one part of the study, participants rated the speaker's confidence, which turned out during piloting to be the most accessible mediation of the term COMMITMENT. In a second part of the study, participants rated the speaker's expectation of a response, which turned out to be the most accessible mediation of the term ENGAGEMENT. The sequence of the two parts of the study was randomly assigned to the participants. The two parts were separated by a cognitive distractor block. I found significant main effects of duration ($\chi^2(2) = 11.256, p = 0.003596$) and pitch excursion ($\chi^2(1) = 74.269, p < 0.001$) for speaker confidence

and a significant main effect of excursion for response expectation ($\chi^2(1) = 85.598, p < 0.001$). Neither part of the study showed interactions of duration and excursion. Hence, shorter duration and a decrease in excursion positively correlated with higher ratings for speaker confidence. Moreover, an increase in excursion positively correlated with higher ratings of response expectation. Figs. 1 and 2 visualize the rating differences for all six types of rising declaratives.

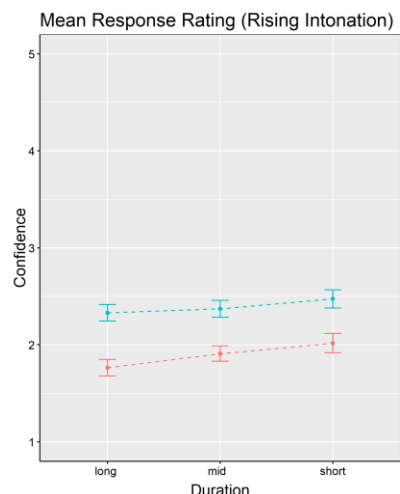


Fig. 1: Confidence ratings

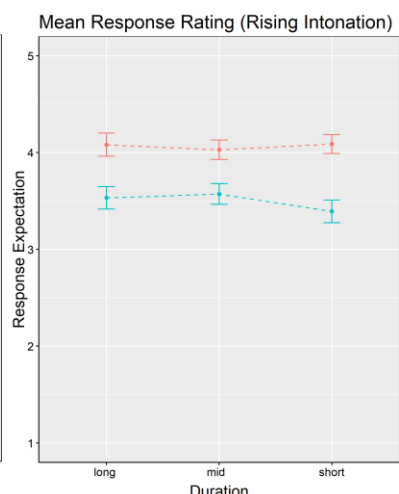


Fig. 2: Expectation ratings

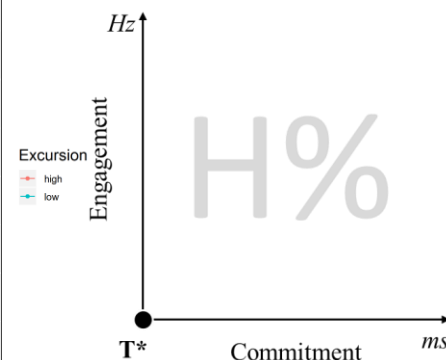


Fig. 3: Prosodic encoding

Discussion & Outlook: The findings reported here support my hypotheses of the interpretability of the shape of the sentence final contour. Changes in pitch excursion and duration independently affected the perception of the speaker's confidence and their expectation of a response, which I consider to be equivalent to the notions of Commitment and Engagement. In response to the challenges these findings pose for bitonal models of intonational meaning, I propose that the temporal and amplitudinal distance from pitch accent to boundary tone can be integrated as an additional measure for capturing the complex encoding of speaker attitudes (see Fig. 3).

References

- Armstrong, M., P. Piccinini, & A. Ritchart. (2015). The phonetics and distribution of non-question rises in two varieties of American English. In *ICPhS*.
- Arvaniti, A. (*In press*). The Autosegmental-Metrical model of intonational phonology. In S. Shattuck Hufnagel & J. Barnes (Eds.), *Prosodic Theory and Practice*. Cambridge, MA: The MIT Press.
- Bartels, C. (1997). *Towards a Compositional Interpretation of English Statement and Question Intonation*, Ph.D. thesis, University of Massachusetts Amherst, Amherst, MA.
- Fletcher, J. & D. Loakes, D. 2010. Interpreting rising intonation in Australian English. *Proceedings of Speech Prosody, Chicago, IL*.
- Ladd, D. R., & R. Morton (1997). The perception of intonational emphasis: continuous or categorical?. *Journal of phonetics*, 25(3), 313-342.
- O'Connor, J. D. & G. F. Arnold. (1973). *Intonation of colloquial English*. Harlow: Longman.
- Pierrehumbert, J., & J. Hirschberg (1990). The meaning of intonational contours in discourse. In: Cohen, P., Morgan, J., Pollack, M. (Eds.), *Intentions in Communication*. MIT P, Cambridge, MA, pp. 271–311.
- Pon-Barry, H. (2008). Prosodic manifestations of confidence and uncertainty in spoken language. In *Ninth Annual Conference of the International Speech Communication Association*.
- Rietveld, A. C. M., & C. Gussenhoven. (1987). Perceived speech rate and intonation. *Journal of Phonetics*, 15(3), 273-285.
- Ritchart, A. & A. Arvaniti. (2014). The form and use of uptalk in Southern Californian English. *Proceedings of Speech Prosody*. Vol. 7. Dublin, Ireland: Trinity College Dublin.
- Tomlinson Jr, J. M. (2009). *Talking it up: The role of temporal context in the interpretation of uptalk*. Santa Cruz, CA: UCSC dissertation.
- Truckenbrodt, H. (2012). Semantics of intonation. In: C. Maienborn, K. von Stechow & P. Portner (eds.), *Semantics. An international handbook of natural language meaning*. Vol. 3. Berlin: de Gruyter, 2039-2969.