

Embarking on PoLaR Explorations: A Framework for Intonational Annotation and Analysis

Byron Ahn,
Nanette Veilleux,
Stefanie Shattuck-Hufnagel,
and Alejna Brugos

draft
November 2022

Contents

Contents	i
List of Figures	ii
List of Tables	ii
1 Conclusions & Summary	1
Bibliography	3

List of Figures

List of Tables

Chapter 1

Conclusions & Summary

PoLaR: A Framework and Toolkit for Advancing Intonational Theory

PoLaR is an annotation framework designed to produce phonologically-informed phonetic labels, and is facilitated by a suite of scripts that aid in labelling and analysis. PoLaR Basic labels capture prosodic structure (prominences and boundaries on the PrStr tier) as well as a set of intonational phonetic characteristics (via f0 turning points in the Points tier, utterance-level or phrase-level f0 ranges in the Ranges tier, and scaled levels relating Points to Ranges via the Levels tier). Advanced labels provide a set of guidelines for more detailed analysis and annotation, such as more finely articulated elements of prosodic structure, association of individual f0 turning points to prosodic structure elements, and more nuanced and interrelated estimates of f0 ranges within and across phrases.

Much recent progress in the domain of intonational theory (in English) has relied on using intonational annotation systems that are phonological in nature. However, these systems were designed to capture contrastive categories that had already been identified, and not the acoustic cues that signal them in different contexts. As a result, what they were not designed to do is capture the full range of meaningful variation in acoustic cue patterns. This situation results in two disadvantages: research on the mapping of acoustic cues to prosodic elements (both within and across language varieties) is hampered; and determining whether the system has identified the correct inventory of phonological categories is unachievable, because a specific inventory of categories must be assumed. Thus, accurate prosodic labelling remains a difficult and time-consuming task for humans, and at the same time cannot be fully implemented by algorithm-based machines.

With PoLaR, we develop a framework that is situated both in acoustic phonetics and the language-specific basics of intonational phonology. In this sense PoLaR produces labels that are methodologically an intermediary bridge (in the sense of being an interface) level of representation (similar to the IPA for binning segmental data). Acoustic cues (which have informed the choice of particular labels in existing phonological labelling systems) are labelled directly, on separate tiers, alongside a set of underspecified phonological labels. In this way, the PoLaR transcription system does not require the labeller to be committed to the precise categories of any intonational phonology (e.g., MAE_ToBI). At the same time, it maintains a clear relationship to the broad categories proposed for in AM models (e.g., pitch accents and boundary tones), while

unbundling the potentially large set of acoustic cues that signal those categories.

PoLaR provides methodological tools for binning together tokens that share a set of acoustic cues and relating them to phonological categories in the prosody. In so doing, it provides a means to investigate facts about the phonetics, the phonology, and the phonetics-phonology interface for intonation. For the same reasons, PoLaR-labelled files can serve as input to machine-learning algorithms, to help reveal which acoustic characteristics are cues that human labellers use to make decisions about phonological categories. Beyond identifying which aspects of the acoustics are seen as relevant, these computer algorithms can be designed and implemented to create a more fully automated (i.e., machine-based) prosodic annotation of both phonetic and phonological labels.

We believe that PoLaR addresses a number of long-standing challenges that have impeded the creation of large prosodically-labelled databases—a resource that has long been desired by the prosodic research community. For example, labelling the acoustic cues alongside basic phonological labels allows more precision in labelling, without requiring the labeller to acquire deep knowledge of (or to routinely ask themselves about) the intonational system as a whole. This is because the system enables quickly-trained individuals to label a single tier (or only one type of event). In addition, we provide explicit guidance about how to interpret and annotate disrupted or distorted f0 tracks (due to, e.g. voice quality and segmental effects), which are historically a bane to both labellers and automatic extraction. Also, certain aspects of PoLaR labelling are facilitated by scripts, such as the automatic labelling of the Levels tier, and the Straight Line Approximation tool for determining the perceptual importance of individual f0 turning points. We anticipate that all of these attributes of PoLaR this will speed up the process of labelling, while also decreasing the amount of anxiety and uncertainty that the labeller experiences. Because the phonetic tiers are designed to transparently track the signal, labellers report feeling more confident and having less to keep in mind. We expect that this will lead to higher rates of inter-labeller agreement, since the task for each labeller is better defined, and the labels themselves are less complicated in their meaning. Moreover, where there is disagreement, the fact that each label is more transparent allows labellers to more precisely target what exactly they disagree upon.

The PoLaR annotation framework is designed to be flexible, to allow straightforward tailoring to new languages and dialects, and also to take into account cues and categories not encompassed by the system as it is described in this monograph. While the focus here is primarily on f0 cues, we foresee extensions to labelling in the PoLaR framework to include other cues such as duration, amplitude and voice quality (see section **?REF?**); the system is flexible enough to incorporate additional tiers and new labels, without requiring a reconsideration of the system as a whole.

The PoLaR labelling system builds on decades of intonational research and is informed by prosodic annotation in many different traditions. We are grateful to the many researchers who have addressed the question of how to annotate prosody, and hope that this set of tools will complement and facilitate prosodic research from diverse theoretical perspectives.

Bibliography