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

List of Figures

List of Tables

Chapter 1

Quick Reference to PoLaR Labels

This appendix serves as a quick-access reference guide, for use while labelling. It provides tables for all Basic PoLaR labels for each tier, as well as the Advanced PoLaR labels. In addition to these tables, some finite state grammar diagrams for the Points Tier labels follow.

Basic Labels

Prosodic Structure (PrStr): a point tier

Label	Phonological Object	Label is time-aligned with ____
*	Prominence	A syllable that has intonational prominence
?*	Possible Prominence	A syllable that might have intonational prominence
]	Phrase's Right Edge	The right edge of the final word of a phrase
?]	Possible Phrase's Right Edge	What might be the right edge of the final word of a phrase

Table 1.1: The Basic labels for the Prosodic Structure tier (for English).

Pitch Points: a point tier

Label	Meaning
0	A turning point in the f0 contour (where line segments meet in a straight-line approximation of the pitch)
?0	Labeller is uncertain whether it is necessary to annotate a turning point in the f0 contour, even after exploring options with the straight line approximation resynthesis tool
0,X	An f0 turning point whose pitch value is approximately X (a labeller's approximation of the f0 when the software-based f0 tracking is unreliable)

Table 1.2: The Basic labels for the Points tier (for English).

Scaled Levels: an automatically-labelled point tier

Label	Given the local pitch range span, the f0 value (in Hz) at corresponding PrStr label time is
1	within the lowest 20%
2	within the 20%–40% range
3	within the 40%–60% range
4	within the 60%–80% range
5	within the highest 20%

Table 1.3: The labels for the Levels tier.

The algorithm used to calculate the the Levels value is based on the currentF0, localMin, and localMax; it is defined below:

```
intervalSize = (localMax-localMin)/5
if currentF0 = f0Max
 then Level = 5
 else Level = 1 + roundDown((currentF0-localMin)/intervalSize)
```

Range Domains: an interval tier

Label	Phonological Object	Example
[*min*]-[*max*]	[*min*] is the local pitch minimum, in Hertz, rounded down to a nearby number ending in a 5 or a 0; [*max*] is the local pitch maximum, in Hertz, rounded up to a nearby number ending in a 5 or a 0	If the local pitch min is 244.2Hz and the local pitch maximum is 381.5Hz, the label should be '`240-385`'

Table 1.4: The Basic labels for the Ranges tier.

Advanced Labels

The rows **highlighted in green** can be considered the "default" labels for this tier; all other beyond this mark some kind of intuition/analysis on the part of the labeller.

Advanced PrStr

Label	Phonological Object	Time-Aligned with
]	Prosodic Phrase Boundary	The end of a Words interval where a prosodic phrase ends
?]	Possible Prosodic Phrase Boundary	The end of a Words interval where a prosodic phrase may end (but the labeller is uncertain)
]]	Large Prosodic Phrase Boundary	The end of a Words interval where a noticeably large prosodic phrase ends
[	Prosodic Phrase Boundary	The beginning of a Words interval where a prosodic phrase begins

Label	Phonological Object	Time-Aligned with
?[	Possible Prosodic Phrase Boundary	The beginning of a Words interval where a prosodic phrase may begin (but the labeller is uncertain)
[[	Large Prosodic Phrase Boundary	The beginning of a Words interval where a noticeably large prosodic phrase

Table 1.5: Advanced PrStr phrasing labels: encoding uncertainty, boundary strength.

Label	Phonological Object	Time-Aligned with
*	Prominence	The center of the Phones/Words interval that is prominent
?*	Possible Prominence	The center of the Phones/Words interval that might be prominent (but the labeller is uncertain)
**	Especially Strong Prominence	The center of the Phones/Words interval that is especially prominent

Table 1.6: Advanced PrStr prominence labels: encoding uncertainty, prominence strength.

Label	Phonological Object	Time-Aligned with
d	Disfluency	A point in the utterance where there is a disfluency
{d	Start of a disfluent region	A point in the utterance where a disfluent region of speech begins
d}	End of a disfluent region	A point in the utterance where a disfluent region of speech ends

Table 1.7: Advanced PrStr labels: encoding disfluency.

Advanced Points

The possible types of phonological object that a turning point is related to, and the pointers to the location of the Prosodic Structure Tier object in time. A special label is used to indicate ambiguity.

Label	Relevant Prosodic Structure object	Relative to an object on the PrStr tier, this f0 turning point is time-aligned...
0	No phonological analysis given by the labeller	N/A
*>	Prominence: *, ?*, or **	...before the relevant *
*<		...after the relevant *
*@		...with the relevant *
]>	Phrase boundary:], ?],]]	...before the relevant]
]<		...after the relevant]

Label	Relevant Prosodic Structure object	Relative to an object on the PrStr tier, this f0 turning point is time-aligned...
]@		...with the relevant]
[>		...before the relevant [
[<	Phrase boundary: [, ?[, [[	...after the relevant [
[@		...with the relevant [

Table 1.8: Advanced Points labels: encoding relationships with the PrStr Tier.

Advanced Ranges

Label	Phonological Object	Example
[*min*]-[*max*]	[*min*] is the local pitch minimum, in Hertz, rounded down to a nearby number ending in a 5 or a 0; [*max*] is the local pitch maximum, in Hertz, rounded up to a nearby number ending in a 5 or a 0	If the local pitch min is 244.2Hz and the local pitch maximum is 381.5Hz, the label should be '240-385'
[*min*]([*min2*]:na)	[*min*] is the Basic ranges value for the minimum, and [*min2*] (written in parentheses) is the Advanced ranges label, capturing labeller intuitions for a value that is not attested locally within that interval	If the local attestable pitch min is 112, but the labeller intuits that the speaker has deliberately not reached a low '110(90:na)-385'
[*max*]([*max2*]:na)	[*max*] is the Basic ranges value for the minimum, and [*max2*] (written in parentheses) is the Advanced ranges label, capturing labeller intuitions for a value that is not attested locally within that interval	If the local attestable pitch min is 192, but the labeller intuits that the speaker has deliberately not reached a high '90-192(300:na)'
X-[*max*]	"X" represents a pitch minimum that the labeller cannot determine; [*max*] is the local pitch maximum, in Hertz, rounded up to a nearby number ending in a 5 or a 0	If the local pitch max is 264.8Hz and the local pitch maximum is not easy to determine, the label should be 'X-270'
[*min*]-X	[*min*] is the local pitch minimum, in Hertz, rounded up to a nearby number ending in a 5 or a 0; "X" represents a pitch maximum that the labeller cannot determine	If the local pitch min is 138.2Hz and the local pitch maximum is not easy to determine, the label should be '135-X'
NA	"NA" indicates a stretch of unreliable pitch tracking, in places where surrounding information cannot be used to infer the pitch range minimum/maximum (*To be used in last resort scenarios*)	If the pitch is a long stretch of unreliable pitch tracking, the label for that unreliable stretch is 'NA'

Table 1.9: Advanced Ranges labels: encoding f0 range information under uncertainty.

Points Tier Finite State Diagrams

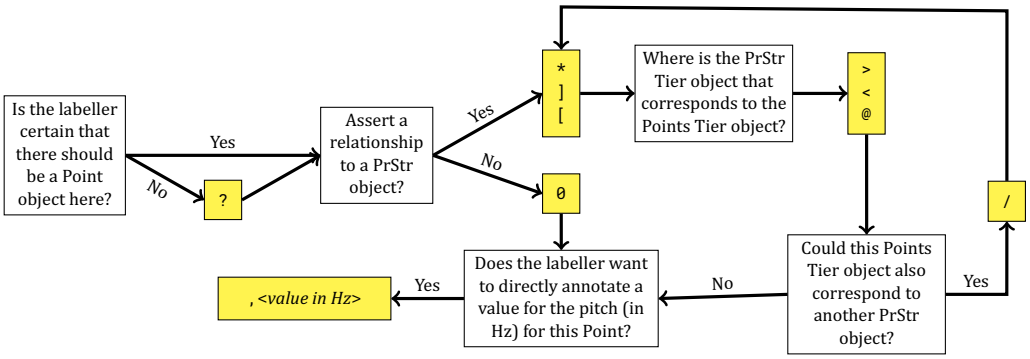

Figure 1.1: A complete finite state grammar for advanced Points tier labels.

Bibliography