

# ADDENDUM: RaP (RHYTHM AND PITCH) LABELING GUIDE, VERSION 1.0

The following are additional conventions for RaP labeling which supplement the RaP labeling manual. Check back regularly for updates. (Issued: 7/24/05)

## I. Additional RaP conventions and clarifications

### Rhythm tier

Within the RaP framework, “perceptual isochrony” is defined as a sequence of metrically prominent syllables or beats, where the temporal intervals formed by each pair of adjacent beats sound approximately equal. Under this definition, a listener might detect some slight speeding or slowing in a syllable sequence which is labeled perceptually isochronous. It is easy to “tap to the beat” in perceptually isochronous speech and labelers will have high confidence regarding which syllables are beats.

When the perception of rhythm is *clear*, label the rhythm that is perceived even if it conflicts with one or more of the conventions. When the perception of rhythm is *not clear*, rely more heavily on the conventions, find the labeling that seems to be most *consistent* with the *conventions* which is not *inconsistent* with your *perceptions* (i.e., violates the fewest of them).

Note that the strength of a metrical beat and the choice between e.g. “X” and “x” is determined with respect to *context*, rather than with respect to some absolute prominence criterion. We have also found that listening to the unfolding information structure helps a lot in determining on which syllables to place X’s. Words which are especially salient informationally will tend to be labeled with X’s.

Given the choice of two equally perceptually salient rhythms, choose to label in the rhythm tier the rhythm which is more consistent with the conventions (usually the faster one).

In very fast speech, the clearest rhythm may occasionally be the “higher level” or slower rhythm. If a rhythm emerges clearly on the first two or three instances of listening from such speech, label the beats for this rhythm, even if they are not 1-2 syllables apart.

When in doubt about the rhythm, prefer a labeling in which beats are left off of polysyllabic function words such as *into* or *under*.

All other things being equal, prefer a rhythm which maximizes the length of the perceptually isochronous sequence.

Note that the label “[x]” should only be used on the lexically unstressed unreduced and reduced syllables of a *polysyllabic* word.

## Tone tier

The rule of thumb for tone labels is “when in doubt, leave it out”. This is particularly relevant for cases in which the relative pitch levels of unstressed syllables with respect to stressed syllables is unclear, so that it is not obvious whether an unstarred tone should be labeled or not. The “leave it out” convention suggests that the marker “+?” is an appropriate label for the unstressed syllable(s) in such cases.

The labels “>” and “<” should be used in cases of expressive use of pitch range where it seems as if the speaker is intentionally aiming for the upper or lower part of his or her pitch range. They can be used even if the speaker doesn’t actually achieve the minimum or maximum pitch that he or she is physically capable of producing. Thus, these symbols are also used to capture instances of a very wide pitch range. Thus, the RaP label sequence <L\* +>H corresponds to the ToBI L\*+H label. Listening to the unfolding of the information structure helps a lot in placing X’s and in placing >’s and <’s.

Moreover, this addendum addresses the issue of ambiguity in the choice of tonal labels. Suppose there is a contour which is *clearly rising* or *clearly falling* over some portion of the speech, but it’s not clear whether an unstressed syllable just before a stressed syllable with a starred tone has the same pitch as the stressed syllable itself. For example, if the contour is falling across a sequence of stressed-unstressed-stressed syllable, then this yields an ambiguity between H\* L+ E\* vs. just H\* L\*. The corresponding ambiguity for a rising contour would be between L\* H+ E\* vs. L\* H\*. The convention adopted to deal with this case is to prefer the following transcriptions for the cases of a falling and a rising pattern, respectively: H\* +? L\* and L\* +? H\*. Transcriptions like H\* L+ E\* and L\* H+ E\* (with the unstarred tone and then an equal starred tone) should not be used unless it’s very clear perceptually and not at all ambiguous that the unstressed syllable truly has the same level as the following stressed syllable. This could be reasonably expected only for speech which is quite slow and careful. (7/24/05)

Another case of tonal ambiguity involves the following. Suppose we have a contour which is primarily level in pitch, such that it’s not clear on exactly which syllable the pitch rises or falls. An example from the RaP manual is <money>. This presents an ambiguity between, say, E+ !H\* vs. just E\* for a case in which it is uncertain on which syllable a rise happens vs. between E+ !L\* and just E\* for a case in which it is uncertain on which syllable a fall happens. The convention adopted to deal with this pattern is to prefer the following transcriptions for these two cases respectively: E+ !H?\* and E+ !L?\*. (7/24/05)

It is important not to confuse RaP’s “bang” symbol (!) used for locally compressed pitch range with the downstep marker used in ToBI.<sup>1</sup> In particular, labelers should be careful to use !L to indicate stepping down in pitch by a small amount, rather than !H as would be used to indicate stepping down in ToBI.

---

<sup>1</sup> “Bang” is a term for the symbol “!” which is commonly used in computer science.

In the case of labeling tones spanning a transition between creaky and modal voice, leave off the bang (!) symbol.

*Placement of labels in textgrids.* For neatness of the visual display in Praat, starred tones should be approximately temporally aligned with the associated “x”, “X?”, or “X” label. Unstarred tones with a “+” to their right should be aligned near the next starred tone on the appropriate syllable. Unstarred tones with a “+” to their left should be aligned near the previous starred tone on the appropriate syllable. Unstarred tones associated with a boundary should be aligned with the appropriate edge of the syllable (near the beginning for utterance- or phrase-initial tones or near the end for utterance- or phrase-final tones).

There will occasionally be situations in which a phrase-final tone clearly seems to be associated with edge of a phrase, even when there is a preceding or following starred tone. In such a case, assign a “+” to the unstarred tone to indicate the relative position of the starred tone, but move the tone label to the appropriate left or right edge of the associated phrase. For example, in <blue\_ties1> the +L on ties is moved to the right edge of the syllable.

## Misc tier

The label “hes” is a sufficient label for flagging cases in which the speaker hesitates but does not produce any additional disfluency.

When using parallelism markers to indicate similarity between non-adjacent tonal sequences, the same number should be used outside parentheses on successive instances of parallel constructions. It does not matter for the purposes of data extraction what the number is or how such numbers are sequentially ordered, merely that similar constructions be assigned the same number. Thus, the following is a hypothetical sequence of labels for parallel constructions placed in the misc tier:

1(// //)1 1(// //)1 2(// //)2 2(// //)2 2(// //)2 3(// //)3 3(// //)3 3(// //)3, etc.

*Placement of labels in textgrids.* Parallelism labels which utilize a parenthesis, including (//, //), 1(//, //)1, etc., should be temporally aligned such that they enclose the associated tonal markers. That is, the label with the open parenthesis should be aligned before the first tone label participating in the parallel construction, while the label with the closed parenthesis should be aligned after the last tone label in the parallel construction

## II. Additional RaP symbols

### Misc tier

//	parallelism between isolated, nonadjacent pitch events, cf. <armani10>
err	speech error

### **III. Updates and Errata to Version 1.0 of the RaP manual**

pp. 35 and 38: The inter-pause interval cutoff criterion for phrase-final tone labeling has been revised downward from 100 to 50 msec.