

Switchboard Prosody Annotation

Jason Brenier and Sasha Calhoun

June 9, 2006

1 Overview

The aim of this work is to annotate a portion of the *Switchboard* corpus with basic prosodic annotation. We will be using, as closely as possible, the ToBI annotation standards for marking pitch accents and phrase boundaries. Changes and additions to these standards are listed below, but before beginning please take time to re-familiarise yourself with the ToBI standards from the ToBI website:

<http://www.ling.ohio-state.edu/tobi>
http://www.ling.ohio-state.edu/research/phonetics/E_ToBI/
http://www.ling.ohio-state.edu/research/phonetics/E_ToBI/singer_tobi.html

2 The Tool

We will be using *Praat* to do this annotation, and will supply you with sound files and TextGrids in Praat format. We are assuming you are already familiar with this program, and have access to a computer where it is installed. If not, please download it and look through the users' manual at:

<http://www.fon.hum.uva.nl/praat/>

3 The Annotation

You should have been supplied with an example dialog, *sw2969B*. Go through the example before starting. Note that you will only hear and annotate one speaker at a time. You will see that there are 6 tiers to the annotation: a Words tier, an Accent tier, a Boundaries Tier, a Breaks Tier, a Notes tier and a Markable Sentences tier. (If these do not appear please contact us before proceeding).

In addition to the changes laid out below, please tell us if there are any changes you think we should make to the scheme, big generalisations we're missing, or any other comments and suggestions.

3.1 Words Tier

The Words tier contains an orthographic transcript of the conversation with a numerical index attached to each word (e.g. *today-20*). You should not make any changes to this tier. Try to keep an unlabelled version of the TextGrid in case you move one of these boundaries accidentally, so you can go back and change it.

3.2 Accents Tier

Mark pitch accents as per the ToBI definition. Do not mark the tonal type of pitch accent, just presence or absence. Mark the pitch accents as close to the target f0 max or min as possible, or in a steady state if there is no max or min. If there is no pitch track during the accented word, or you can't see it clearly, try zooming in and going to Pitch >> Advanced Pitch Settings and adjusting the range and voicing threshhold. If there is still no pitch track where the peak appears to be, mark the peak at the latest point before the pitch track ends and then make a note in the Notes tier 'no pitch track at peak'.

Mark each accent with the numerical index of the word it is associated with, e.g. if *today-20* has a pitch accent, you would mark this *20*. If the peak is before or after the word boundary, add a 'early peak' or 'late peak' marker in the Notes tier. If there are two accents in one word, which is rare but sometimes happens, mark the first accent, e.g. *20-a* and the second one, e.g. *20-b*. There should be at least one accent per phrase (3 or 4 boundary).

Each accent should be marked according to its *strength*:

- **Nuclear (N)** The most prominent accent in the phrase, the nuclear accent, is marked with N, e.g. *20N*. We are interested in *structural* and not *phonetic* prominence, so listen for which word you think the speaker makes sound most important. This will normally be the right-most accent, unless the last accent is heavily reduced, i.e. sounds like an afterthought or something that's already known. If in doubt, try to frame a question the speaker sounds like they're answering, the answer is the N.
- **Strong Pre-Nuclear (PN)** If there is an early high, *acoustically* prominent accent followed by a later lower accent which is nevertheless structurally important to the phrase, then the first accent is PN, the later N. For example, see 166.3-167.9 in sw2969B, *well YOU_{PN} should have just ripped the whole THING_N out*.
- **Regular Accent** Regular pre- and post- nuclear accents are marked as noted above with the index of the word they are associated with.

- **Question Accents (?)** If an accent sounds weak but there is a clear pitch movement; or there is a clear prominence, but no pitch movement; and it is not nuclear, mark it as a ? accent, e.g. *20?*
- **No Accent in Phrase (Z)** If there is a short phrase, particularly a filled pause, disfluency or floater, where none of the words sound accented, mark one of the words in the phrase with a Z, e.g. *20Z*.

3.3 Boundaries Tier

Mark phrase boundaries with a break index strength of 3 or 4, and give each phrase the appropriate label '3' or '4'. If you need refreshing on the difference:

http://www.ling.ohio-state.edu/research/phonetics/E_ToBI/ToBI/ToBI.3.html

You do not need to mark phrase accents or boundary tones. NB - Phrase boundaries should align exactly with the word boundaries in the Words tier, so we can align words and phrases in our database.

3.4 Breaks Tier

In addition to break strength, disfluent and incomplete phrases should be marked in the Break Tier using the following two labels:

- **Disfluent Phrase (2p)** This corresponds to ToBI 1p, 2p and 3p and should be used to mark:
 - cut-off phrases, particularly before restarts or repetitions
 - hesitations mid-sentence, where the speaker sounds like they don't know what to say next, **not** that they are pausing for effect, e.g. to highlight the next word
 - floaters: where a phrase sounds fluent, but then there's a extra one or two disfluent words at the end. Put a regular phrase break after the last fluent word, then put a 2p break after the floater
- **Backchannel Phrase (X)** For short non-disfluent phrase that don't sound like a full intonation phrase, usually things like *um*, *yeah*, *right*, *you know*.

3.5 Notes Tier

Please mark anything noteworthy in your annotation or the orthographic transcription. Particularly:

- If there is a transcription error, use S, I and D to indicate the substitution, insertion or deletion of a word, e.g. *D-dog*. NB - if an accented word has been deleted, mark this as, e.g. *382-1* in the Accent tier (if it occurs immediately after word 328), and also note the deleted word in the Notes tier. The purpose of this annotation is not to fix the transcription, so only note these errors if you think the error makes your annotation unreliable, e.g. accents on deleted words or missing words at phrase boundaries
- Very occasionally the transcription is so bad it makes prosody annotation near impossible, i.e. when there are 5-6 words together all misaligned. In these cases mark the beginning and end of the region as ‘untranscribable region begins’ and ‘untranscribable region ends’ and do not annotate within this
- Mark contractions that are not transcribed as such using, e.g. *contraction-wanna*
- If the word boundaries are badly off, note this as, e.g. *misalign-dog*
- As noted above, mark of there is a late or early peak (outside the word boundary); or if there is no pitch track at the apparent *f0* peak
- If there is a word that sneaks in to the end of an intonational phrase (that is, the phrase almost seems to end just before the word, but the word sounds like it belongs in the phrase anyway and not in its own phrase), label it as a *floater* in the notes tier. However, put a phrase break before the floater and call the break after the floater 2p as noted above.
- As noted below, if the beginning or end of the markable sentence is misaligned (marked with *), move the boundary and note ‘markable boundary moved’

If there are any other phenomena which you find coming up often, let us know and we will make a special label for it.

3.6 Markable Sentences

This tier indicates the sentences that we are interested in, so you only need to annotate the speech between these markers. All other sentences can be ignored. Note that this tier was produced using a slightly different transcript to the one you’re using, so the transcript of the sentence might be different - don’t worry about this. However, sometimes we couldn’t determine the time of the beginning or end of the markable sentence. In these cases, the first or last word is marked with a *. If this happens, move the boundary of the markable sentence so that it aligns with the actual beginning of the sentence in the Words tier and note this in the Notes tier. Annotate the whole resulting sentence. An example of this can be seen at *sw2969B*, 0:08.7:

(Markable Sentence) : s9_500 *Well the topic today was gardening and lawn care

(Words Tier) : I heard that the topic today was gardening and lawn care

You shouldn’t have to move the boundary more than 2-3 words.