

## 2.10 Break index 2

*Note for this 2006 ToBI tutorial:*

*There is one remaining break index that we have not yet discussed: break index 2. The 1993 ToBI tutorial (Ohio State) noted that “In the course of developing the ToBI transcription system, we encountered several utterances in which we felt a strong sense of disjuncture at a boundary between two words where the pitch pattern showed no evidence of the necessary tonal events for either of these (3 or 4 break level) levels of intonational constituency,” As described in detail below, break index 2 was devised to label these mismatched cues that provide conflicting evidence either for an intermediate phrase break or an ordinary intra-word break. However, since then it has not proved to be a simple matter to divide the set of potential 2’s into the two kinds of mismatch described below in the 1993 ToBI Labelling Guide. Cases like ‘Six southern Iraqi cities’ in the discussion below, where timing cues suggest 3-breaks but there are no compelling intonational cues to a Phrase Tone, are not unusual, but cases like ‘Quincy’, where timing cues are consistent with an ordinary 1-break but intonational cues suggest a Phrase Tone, are often parsed differently by different transcribers. In addition, some labelers find the 2 diacritic useful for cases in which there is no actual mismatch, but the labeler perceives a boundary between two words that is larger than a 1 but smaller than a 3. We are currently discussing options for re-phrasing the discussion of Break Index 2, and it is anticipated that changes may be made to this section to reflect evolving understanding of this level of disjuncture. For now, we include the current sections of the 1997 3<sup>rd</sup> revision of the ToBI Labelling Guide.*

*From the 1997 3<sup>rd</sup> revision of the ToBI Labelling Guide*

([http://www.ling.ohio-state.edu/~tobi/ame\\_tobi/labelling\\_guide\\_v3.pdf](http://www.ling.ohio-state.edu/~tobi/ame_tobi/labelling_guide_v3.pdf)):

As noted in the previous section, each 3 on the break index tier must correspond to the marking of a phrase accent for the intermediate phrase on the tone tier, and each 4 must correspond to the marking of a boundary tone. The implication is that any other interword juncture will be something that can be transcribed on the break index tier with either a 0 or a 1. However, the subjective impression of boundary strength does not always allow such a neat correspondence. In the course of developing the ToBI transcription system, we encountered several utterances in which we felt a strong sense of disjuncture at a boundary between two words where the pitch pattern showed no evidence of the necessary tonal events for either of these two levels of intonational constituency. We also encountered the converse case: utterances in which the pitch pattern at a boundary between two words clearly indicated an intermediate or intonation phrase boundary with none of the preboundary lengthening or other cues that support the subjective sense of a strong disjuncture. Break index 2 was devised to mark cases of these two types of ‘mismatch’ between the subjective boundary strength and the intonational constituency. These two types are described in the ToBI Annotation Conventions as follows:

a strong disjuncture marked by a pause or virtual pause, but with no tonal marks; i.e. a well-formed tune continues across the juncture.

OR

a disjuncture that is weaker than expected at what is tonally a clear intermediate or full intonation phrase boundary.

Example utterance <<iraqi>> illustrates the first type of mismatch, and example utterance <<quincy>> illustrates the second. In <<iraqi>>, the smooth sequence of apparent downstepped peak accents with no clear intervening phrase accent suggests that the words "six", southern, "iraqi", and "cities" all belong to the same intermediate phrase, yet there is an intonation phrase sized pausing between each adjacent pair of these words. In <<quincy>>, the clear tonal markings for at least an intermediate phrase boundary are unaccompanied by any clear preboundary lengthening, making some transcribers uncomfortable in labelling this juncture with a 3.

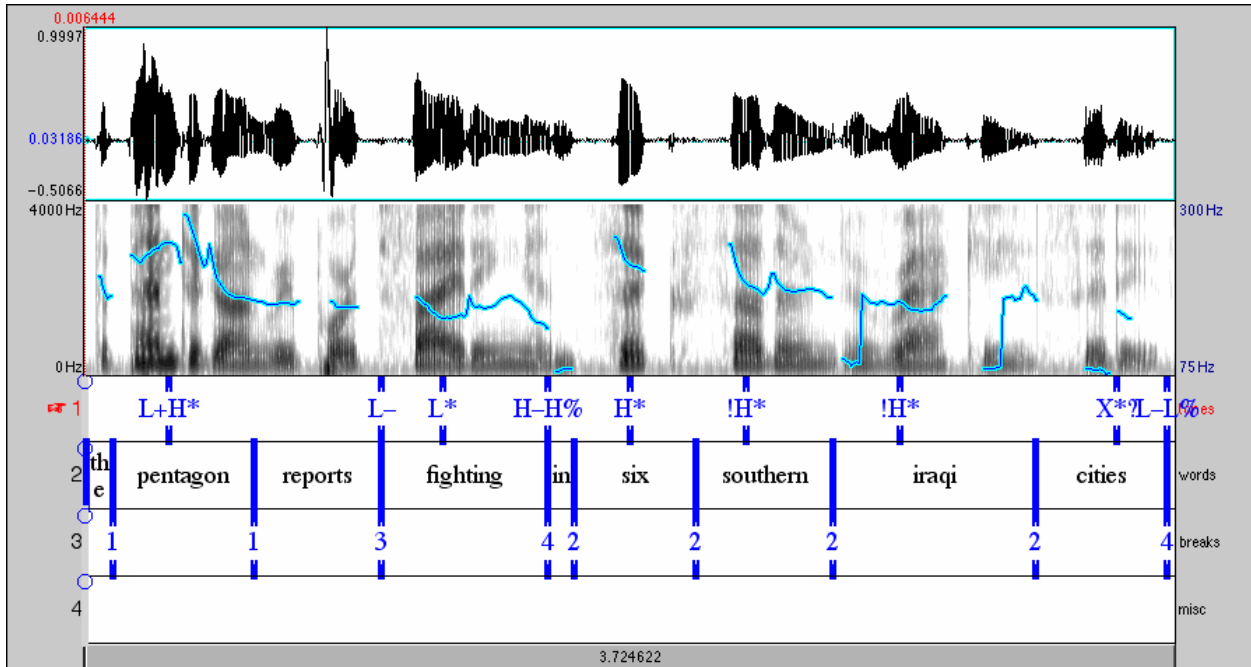


Figure 2.10.1

<iraqi>

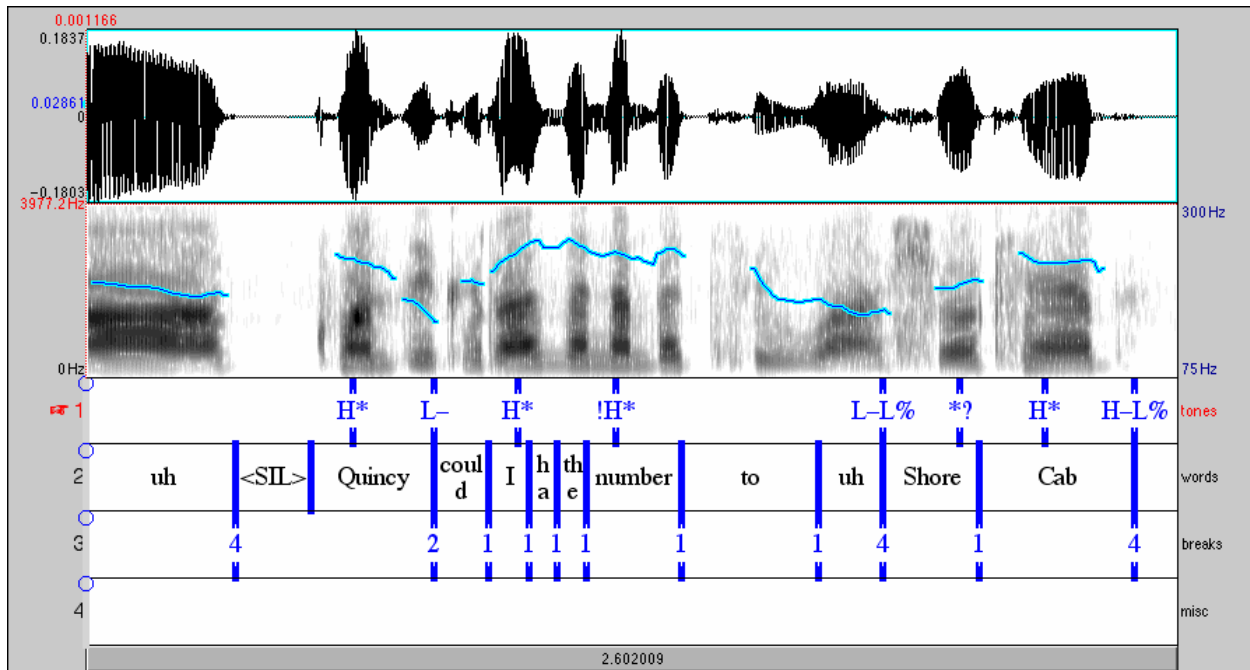


Figure 2.10.2

<quincy>

Break index 2 was devised for cases where the mismatch between the tonal marking and the disjuncture is not accompanied by any sense of hesitancy or disfluency. When 2 is used in the first way (to indicate a stronger sense of disjuncture than 1 even while producing a coherent contour for an uninterrupted intermediate phrase), it can have the rhetorical effect of careful deliberation, as in the <<iraqi>> example. In the opposite case (when 2 is used to mark intermediate phrase boundaries which do not have a very strong sense of disjuncture) the speaker may be speaking quickly to hold the floor or to convey a sense of urgency, while using the tonal marks necessary to convey attentional focus on several closely placed words. We suspect that both types of 2 will be explained ultimately by a better understanding of the complexities of discourse structure, an understanding that can best be achieved by the transcription and analysis of many occurrences in natural dialogue.

### Summary of ToBI labels introduced so far:

Tones:

H\*: high pitch accent

L\*: low pitch accent

L+H\*: bitonal pitch accent with low tone followed by high tone prominence

L\*+H: bitonal pitch accent with low tone prominence followed by high tone

!H\*: downstepped high pitch accent

L+!H\*: bitonal pitch accent with low tone followed by a downstepped high tone prominence

L\*+!H: bitonal pitch accent with low tone prominence followed by downstepped high tone

H+!H\*: bitonal pitch accent with high tone followed by downstepped high prominence

L-L%: low phrase tone, low boundary tone

H-H%: high phrase tone, high boundary tone

L-H%: low phrase tone, high boundary tone

H-L%: high phrase tone, low boundary tone

H-: high phrase tone  
L-: low phrase tone

Break indices:

- 0: word boundary erased
- 1: typical inter-word disjuncture within a phrase
- 2: mismatched inter-word disjuncture within a phrase
- 3: end of an intermediate phrase
- 4: end of an intonational phrase