**2.11 The p diacritic (and the %r tone label)**


 *Note for this 2006 ToBI tutorial:   Prosodic labelling of disfluencies has turned out to be one of the biggest challenges to the ToBI system, and because these events happen often in spontaneous speech, extending the system to label them appropriately and reliably is a pre-requisite for labeling large spontaneous corpora of spontaneous speech.  This is a topic of current research (e.g. Arbisi-Kelm and Jun (2005)[1]) It is anticipated that changes may be made to this section to reflect evolving understanding of the prosody of disfluencies. For now, we include the section that describes the  p diacritic for disfluent regions, and the %r tone label for a restart from the 1997 ToBI Labelling Guide.*

*From the 1997 3rd revision of the ToBI Labelling Guide*
(http://www.ling.ohio-state.edu/~tobi/ame_tobi/labelling_guide_v3.pdf):

[Christine Nakatani and Elizabeth Shriberg contributed greatly to the preparation of this and the following sections.]

There are other cases of mismatch between tone tier and segmental rhythm, however, where break index 2 does not seem to be appropriate. For example, in utterance <<display>>, the pauses after "Baltimore", "which", and "leave" do not have the feel of a speaker striving for an effect of judicious deliberation, as in the "six southern Iraqi cities" phrase of the <<iraqi>> example, but rather sound disfluent, as if the speaker were hesitating as he searches for the next word. Such cases can be distinguished from fluent cases of 2 by the use of the p diacritic.
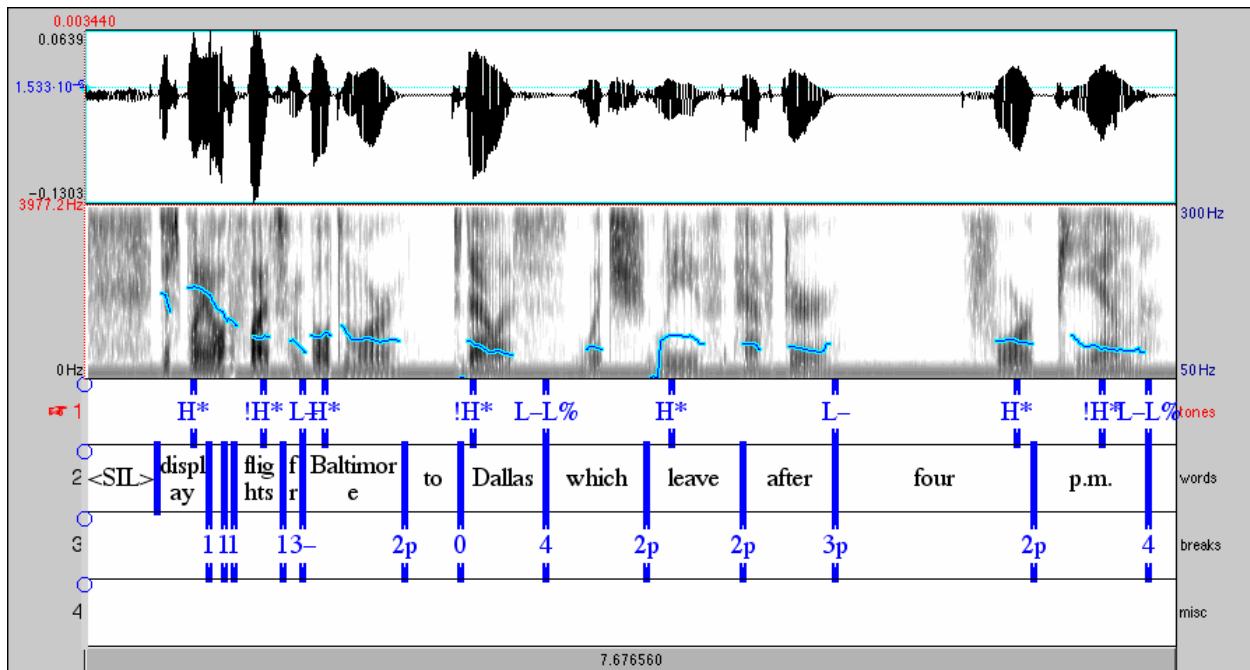


EXAMPLE <display>: Display all the flights from Baltimore to Dallas

_____

[1] Timothy Arbisi-Kelm and Sun-Ah Jun (2005) "A comparison of disfluency patterns in normal and stuttered speech" in *Proceedings of Disfluency in Spontaneous Speech Workshop*, Aix-en Provence, France.

```
       1  1  1    1  3-     2p 0    4
     which leave after 4:00 p.m.
        2p   2p   3p  2p  4
```

The p diacritic is used in conjunction with a break index 1, 2, or 3, to indicate a disfluency in the timing or separation of words across a break.  The notation `p' was chosen initially to denote the prolongation of the hesitation pause with break indices 2 and 3, but we have since extended the diacritic's usage to cover also abrupt cutoffs before restarts and repairs, which are often but not necessarily separated from the disfluent stop by a pause.  In this case, the appropriate break index is 1.  Thus the inventory of combinations of break index and p diacritic is:
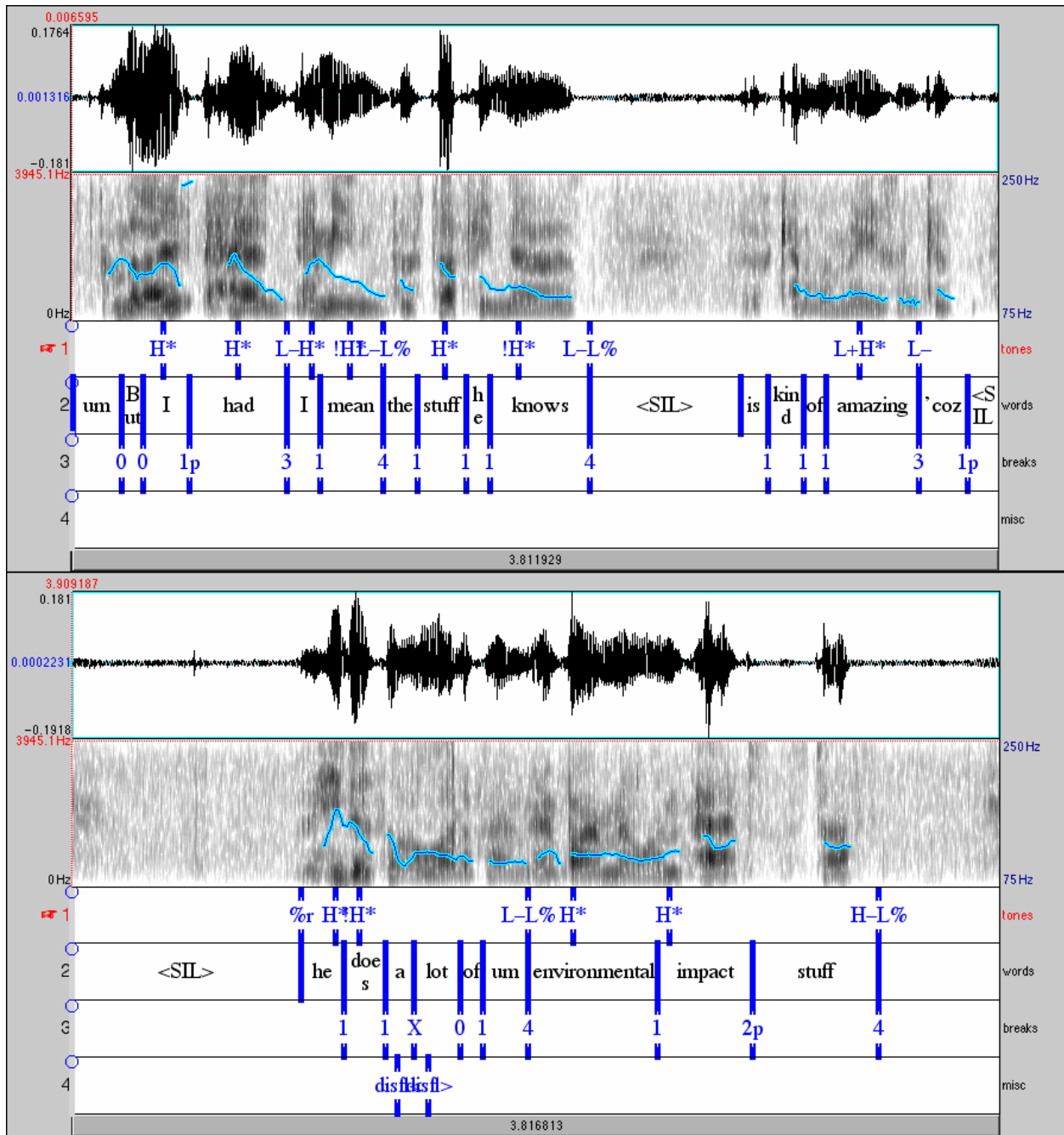
  1p -- an abrupt cutoff before an actual repair, or as if stopping to permit a repair or restart of some kind
  2p -- a hesitation pause or prolongation of segmental material where there is no phrase accent perceived in the intonation contour
  3p -- a hesitation pause or a pause-like prolongation where there is a phrase accent in the tone tier.

The p diacritic is not used with break index 4, because it is difficult to reliably identify hesitations between two full intonational phrases.  Example utterances <<amazing>> and <<cheapest>> illustrate the use of the diacritic with break indices 1 and 3. Example <<display>> also had an occurrence of 3p.  Note the presence of the phrase accent distinguishing this interword juncture from the surrounding cases of 2p.

```
EXAMPLE <amazing>:       um But I had I mean the stuff he knows is kind of
                 0  0 1p 3 1   4  1   1 1   4 1   1 1
              amazing 'coz he does a lot of um environmental
                  3   1p 1   1 X  0  1 4          1
              impact stuff
                 2p   4
```

First panel — tones: H*  H*  L–H*  !H* L–L%  H*  !H*  L–L%  L+H*  L–

words: um | B ut | I | had | I | mean | the | stuff | h e | knows | <SIL> | is | kin d | of | amazing | 'coz | <S IL>

breaks: 0  0  1p    3  1    4  1  1 1    4    1  1 1    3  1p

Second panel — tones: %r  H*H*  L–L% H*  H*  H–L%

words: <SIL> | he | doe s | a | lot | of | um | environmental | impact | stuff

breaks: 1  1  X  0  1  4    1    2p    4

misc: disfldisfl>

EXAMPLE <cheapest>: I want to see the cheapest flight from Atlanta
           1   1 1 1 3p    1   1  1   3

         to Baltimore
          1     4

0.011073
0.1579
−0.001204
−0.1265
3977.2 Hz

250 Hz

0 Hz                                                                75 Hz

| | H* | H* | L– | H* | !H* | L* | H– | L* | H–L% | tones |
|---|---|---|---|---|---|---|---|---|---|---|
| I | want | to | see | the | cheapest | flight | from | Atlanta | to | Baltimore | words |
| 1 | 1 | 1 | 1 | 3p | 1 | 1 | 1 | 3 | 1 | 4 | breaks |

misc

3.672114
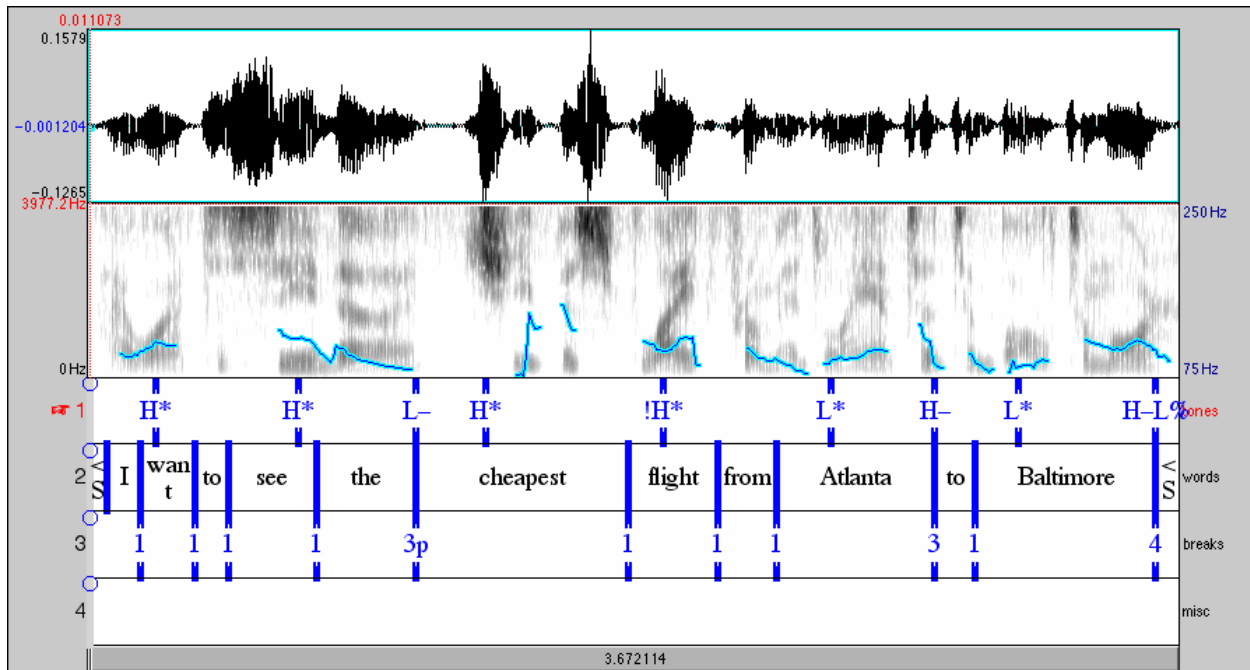
In general the p diacritic should be used conservatively, and should not become a substitute for 2. A good test for appropriateness is to imagine whether the break would have been the same if the speaker were asked to repeat the utterance with the same intonation, but more `fluently'. If the break were the same upon repetition, it should probably not get the p.

Note also that the prolongation of segmental material for a 2p label can physically occur at the beginning of a word rather than at the end, as in example <<least>>, where the hesitation lengthens the [l] of "least" rather than the vowel of "the".

EXAMPLE <least>: Between Boston and Denver I'd like to a flight that
                    3    1  1    4  1  1 3p 1   1   1
            takes the least amount of stops to get to Boston
                3p 2p   1    0 1   4 1  1 0   4
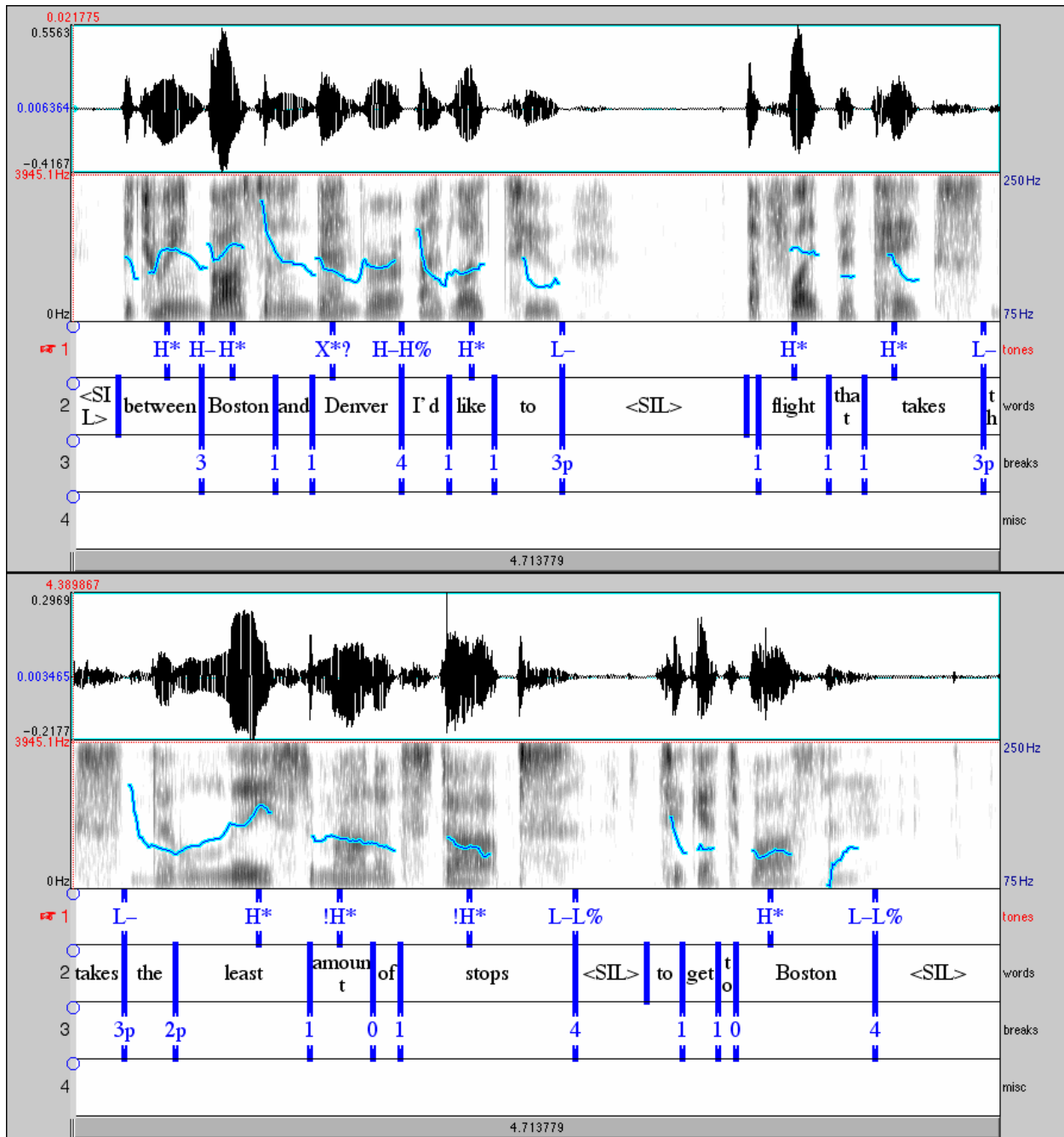
*Comment (7/7/05): One point to keep in mind is that the 'cutoff' described for 1p is not equivalent to the loss of phonological material from the end of a word. For example, speakers often apparently delete the final stop from an consonant cluster, as in 'los' an' found' for 'lost and found'. In these cases there is clearly no disfluency and 1p should not be used. Conversely, an abrupt cutoff that needs a 1p can occur when no phonological material is missing, as when a final vowel is abruptly interrupted, e.g. in an utterance like "I-I don't know". In these cases there is a perceived disfluency i.e. a sudden and unnatural interruption (often at the glottis), despite the fact that the final segment of the word was produced, and a 1p would be appropriate. Perhaps the best definition of a cutoff that requires a 1p label is that the word ends in a way that*

*sounds interrupted.*


*Here is what the labeling guide says about %r:*

Closely associated with these definitions of 1p, 2p, and 3p in the break index tier is the tone tier label %r, for restarting with a brand new intonation contour when the last contour was interrupted without being finished by some disfluency.  This is most common at a `repair', where the speaker abruptly stops and begins again with the intended or `repaired' material, as in example utterance <<amazing>>, already cited above, and in example <<connections>>, below.

```
EXAMPLE <amazing>:       um But I had   I mean     the stuff he knows
                    0   0 1p  3   1    4    1   1 1    4
                      H* H* L- H* !H* L-L%      H*     !H* L-L%
                is kind of amazing 'coz    he does a lot of um
                  1   1 1      3  1p  1    1 X   0  1  4
                      L+H*  L-     %r H* !H*          L-L%
                environmental impact stuff
                        1     2p   4
                   H*        H*        H-L%
```
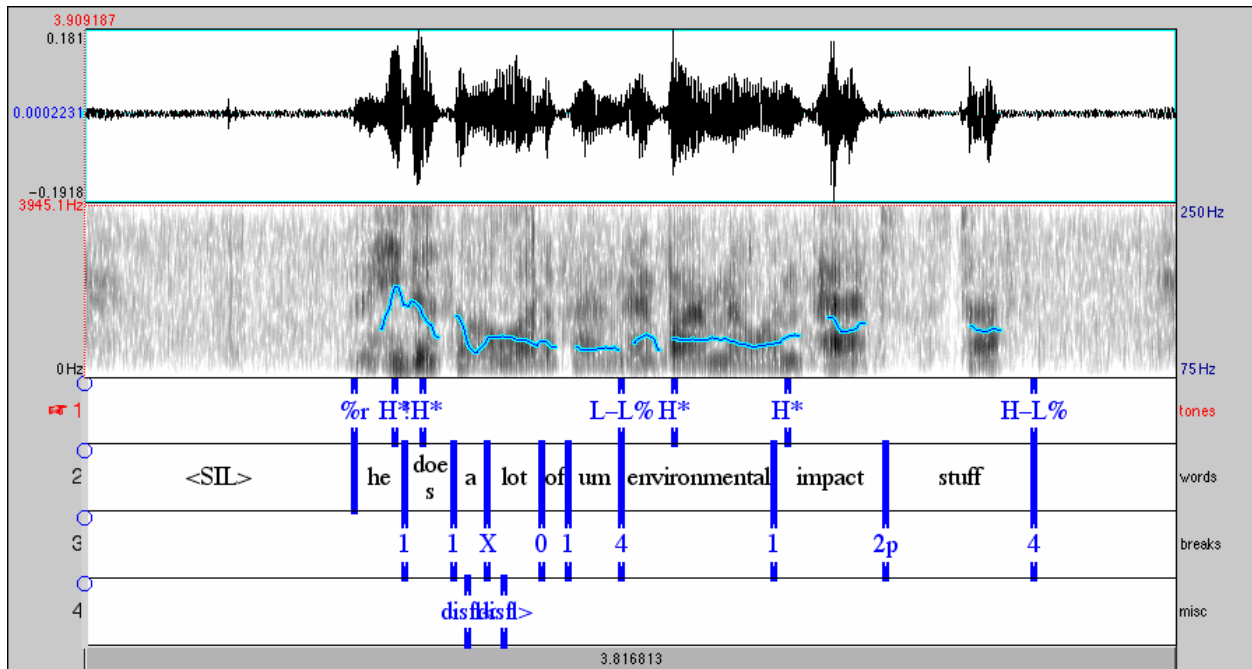


```
EXAMPLE <connections>:  What are the plane sizes   for these flights and
                    1  1  1    1    4   1    1     4 1
                  H*        L*      H-H%    H*   H* H-L%
                do they ha(ve)- do  are there any other flights
                  2p   1     1p  1p  1    1   1    1     1
                            %r         H*        !H*
                that have s-    connections
```
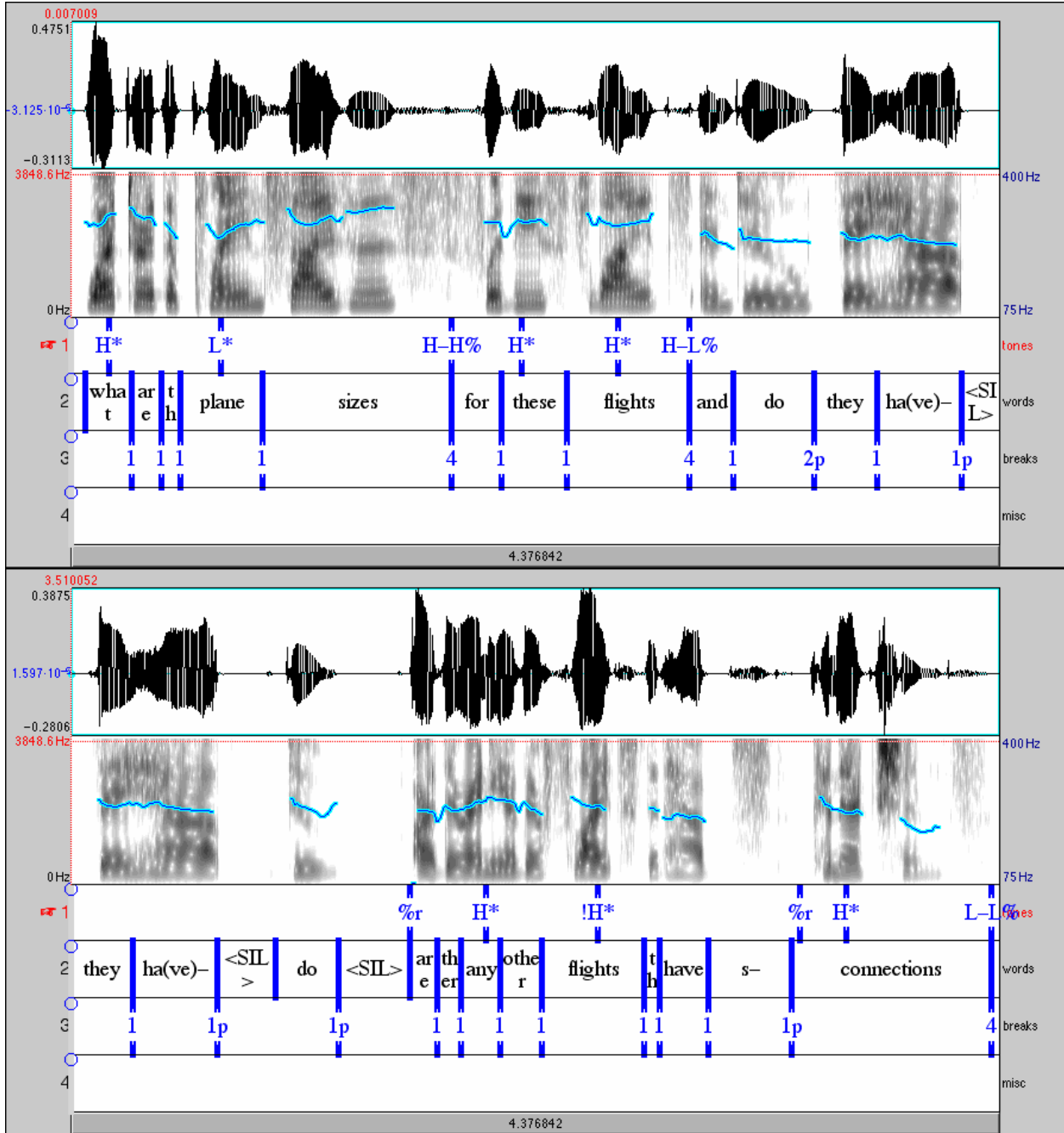
As with the use of the p diacritic, one should be conservative in using %r. It is needed only if there is good evidence that a new intonational phrase has begun after disfluent pause, evidence such as a notable change in f0 range or amplitude. It should not be used in cases such as the "had" after the first 1p in <<amazing>>, which continues with a fluent H* accent in the same pitch range (unlike the H* !H* on "he does" after the second 1p in this utterance, which is in a new pitch range). Nor should %r be used in example utterance <<abbreviation>>, where after

the speaker stumbles and pauses momentarily around the end of "what is the", the intonation on "abbreviation" continues as if there had been no interruption.

EXAMPLE <abbreviation>: What is the b-  abbreviation   n          under
                 0  1  1  1p          3-    3    3p
                      H*      H- L+H* L- H* !H-
          the category d c  mean
           1      1 1  1   4
          H*     H* H*    L-L%
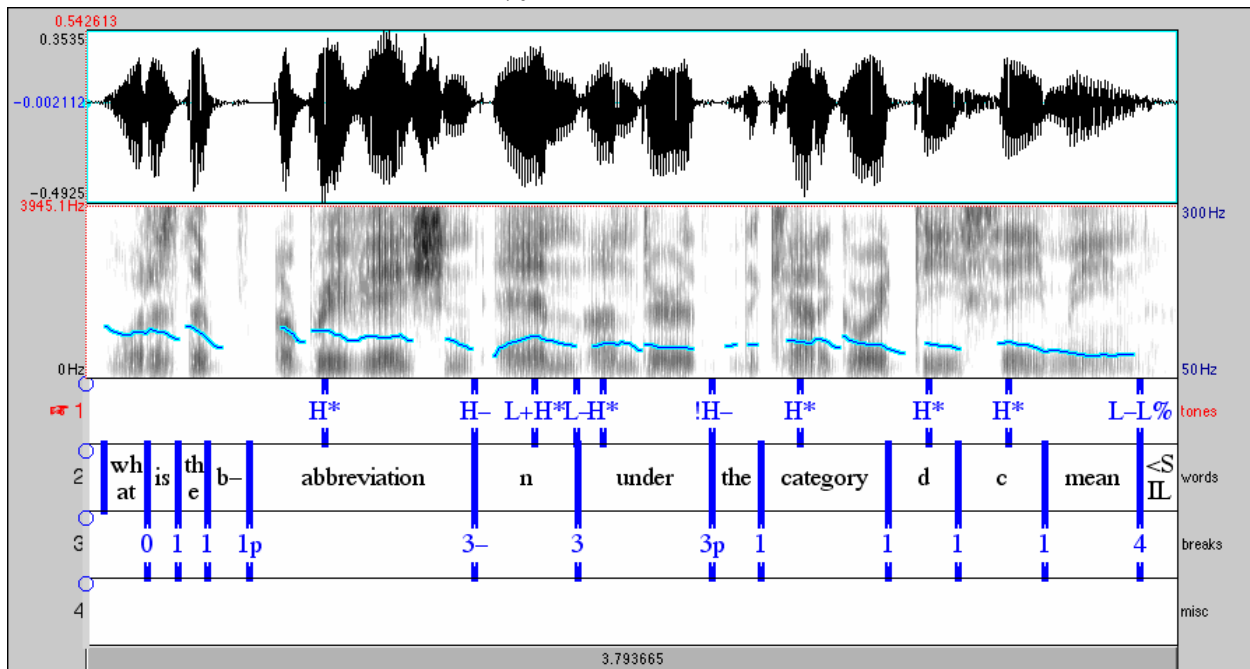


Especially, %r should not be used after a 3p, where the (re)start of a new intonation contour is already implicit in the break index for the intermediate phrase.

**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

Disfluency marked break indices:
1p: an abrupt cutoff before an actual repair, or as if stopping to permit a repair or restart of some kind
2p: a hesitation pause or prolongation of segmental material where there is no phrase accent perceived in the intonation contour
3p: a hesitation pause or a pause-like prolongation where there is a phrase accent in the tone tier.

Other disfluency markers:
%r: marks the well formed beginning (restart) of an intonational phrase after a disfluent ending of the previous intonational phrase

Optional labels:
<: late High Tonal peak