2.12 Uncertainty in labelling

While many stretches of speech will be straightforward to label with ToBI, many others may prove more challenging to the labeller. Naturally-produced speech contains an enormous amount of variation, and there are cases where a labeller may be uncertain what labels to use. There are times when the interpretation of the prosody will be ambiguous, and more than one ToBI label may seem plausible.

Table 2.12.1 sources of variation that may lead to uncertainty

• individual differences between speakers (obvious differences like pitch, high-pitched or low-pitched voice)
• differences even for the same speaker (changes in rate, loudness, affective differences—boredom vs. excitement)
• dialectal differences: speakers from different parts of the country, while still speaking what is considered Mainstream American English, may pronounce some words differently. Likewise, there may be some dialectal variation in intonation
• stylistic differences (news broadcasters or lecturers vs casual conversation)
• a host of other confounding factors (which will be addressed in a new chapter in a future version of this tutorial): speech rate, segmental effects, ambient noise if the recording was not done under ideal conditions like static or background noise, signal dropout, other speakers, laughter, yawning, coughing, irregularities of voicing due to illness (laryngitis or a cold) or even tiredness.

The ToBI system provides a variety of conventions for addressing ambiguity and other sources of labeller uncertainty, including uncertainty labels and the alternatives tier.

2.12.1 The alternatives (alt) tier *(Note for 2006 draft: all content on the alt tier is still under development, and the conventions described are subject to change)*

The alt tier, a recent addition to the ToBI system, allows the labeller to indicate places where more than one label was seriously considered. There may be times when a labeller spends a comparatively long time determining which of two possible labels to use for a particular tone or break. The alternatives tier allows an outlet for the labeller to include both labels, so that she can move forward with more speed, and less frustration.

Use of the alternatives tier, along with other markers of uncertainty which will be discussed later in this section, also provides a record of which regions and points in time in the speech signal are harder to label, so that researchers may investigate the sources of ambiguity. Without uncertainty markers, labels would not differentiate between those stretches of speech which are straightforward to label, and those which contain ambiguities.

2.11.2 Uncertainty regarding pitch accent type

A common situation of uncertainy arises when the labeller has determined that a syllable is pitch-accented, but has trouble deciding which pitch accent type best accounts for the pitch pattern observed. Most often, the labeller is considering between two specific pitch accent types, such as H* vs. L+H*. The labeller might feel that the best label to use would be H*, but that L+H* is a reasonable alternative. In such a case, the labeller should use the label H*? in the tones tier, and put L+H* in the alt tier. The ? symbol indicates, simultaneously, that the labeller was somewhat uncertain about the choice of pitch accent label, and that the alternatives tier has been used to indicate the other pitch accent that the labeller considered. The example <indeed>, shown in Figures 2.12.1 & 2.12.2, contains two instances where the labeller was not certain about which pitch accent label to use: first on the syllable –*deed* of *indeed*, and then later in the file, on the monosyllabic word *clouds*. In the first full intonational phrase of the file (shown in Figure 2.12.1) with the words *and indeed*, there is a strong prominence that is marked by a peak in the pitch track on the syllable –*deed* of *indeed*; the syllable is clearly pitch-accented. However, the labeller was not certain whether the pitch accent on –*deed* should be labelled with the single-tone H*, or with the bitonal L+H*. The pitch at the beginning of the phrase starts out moderately high in the speaker's pitch range, and rises slightly into the first syllable of *indeed*, –*in,* and then rises to a peak in the pitch-accented syllable. There is an apparent fall in f0 shown in the pitch track at the end of the syllable *in-*, but this can be attributed to segmental effects. In fact, these segmental effects between the two syllables of the word *indeed* make the f0 tricky to interpret. The rise into the pitch accented syllable appears to be quite sharp, as would be expected from the bitonal pitch accent L+H*. However, the labeller here felt that the pitch accent sounded more like an exaggerated or emphatic production of a single-tone H* pitch accent.
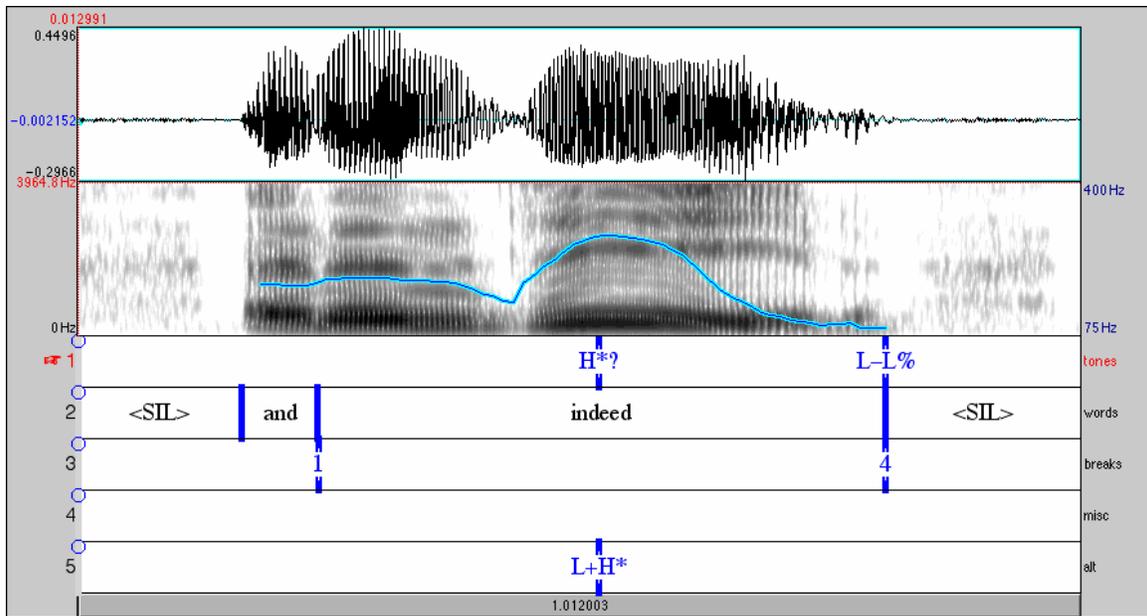
Figure 2.12.1 Using the alt tier and the ? symbol to indicate uncertainty about the choice of pitch accent label in the first intonational phrase of the file indeed.wav   <indeed>

In the second Intonational Phrase of the file <indeed>, shown in full in Figure 2.12.2, below, the labeller has used L+H*? on the word *clouds*. In this case, the second option considered by the labeller was L+!H* (where the bitonal has a downstepped High). In this case, the labeller perceived the pitch height of the peak of the pitch-accented word *clouds* as being about the same as the pitch height reached in the preceding pitch accent on the *–lec–* of *electron*. However, the pitch track shows that the f0 of the peak on *clouds* as lower than the peak of the first pitch accent. The labeller gave precedence to the choice of L+H* (with no downstep), and put that label in the tones tier, along with the "?" marker. Then, she put her second choice, L+!H*, in the alt tier.
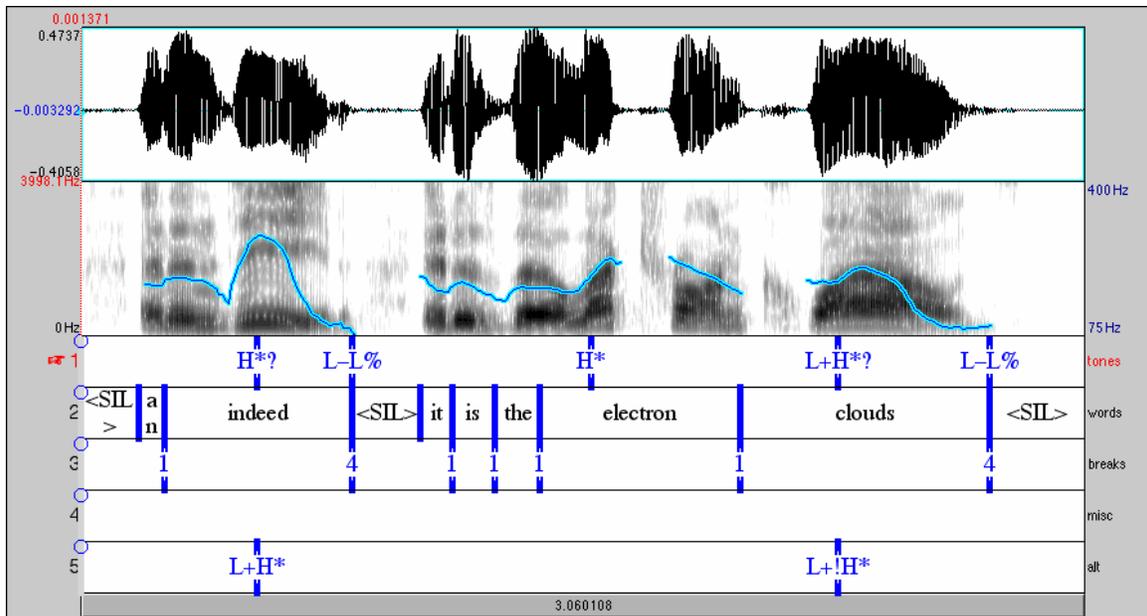
Figure 2.12.2 Another example of using the alt tier on *clouds* in the second intonational phrase   <indeed>

2.12.3 The X*? label:

While a labeller will usually have a first choice of label when considering two alternatives, there may be times when the labeller will be unable to decide which pitch accent type to put in the tones tier. The label X*? can be used for such cases. Whenever possible, the labeller should still use the alternatives tier, and list both alternatives there. For example, if the labeller can't decide between L+H* and H*, and isn't happy committing to either one for the tones tier, the labeller can put X*? in the tones tier, and both L+H* and H* in the alt tier. The labeller should align a single point in the alt tier with the X*? in the tones tier, and list both alternatives at that point, separated by a "pipe" ("|") symbol, eg. "H*|L+H*".

Table 2.12.2 Commonly confusable pairs of pitch accents

H* vs L+H*:   The pitch accented syllable clearly has a High prominence, but the rise may be more gradual than expected for the bitonal L+H*, but steeper than would be expected for a single-tone H*, and there may be varying degrees of evidence for a preceeding L target.
L* vs H* (in a compressed pitch range):   The syllable is clearly prominent, but the pitch track is more or less flat, with no clear peaks (as often occurs with H*) or dips (as one would expect for L*), and the labeller is uncertain whether the region in question is low enough in the speaker's pitch range to justify Low tones (L*) or in a mid-range, such as would happen with H*.
L+H* vs L*+H: There is a distinct low and a distinct high associated with a pitch accent, but the labeller is uncertain which tone, L or H is more strongly associated with the prominence
H* vs !H*:   A pitch accent following another High tone pitch accent (such as H* or L+H*) is

somewhat lower in f0 than that preceding accent, but the labeller is uncertain that the tone is perceptually low enough to be a !H*.

H+!H* vs !H*:
The labeller is uncertain whether a peak leading to the !H* prominence is associated with that pitch accent, or if it is associated with some preceding High tone pitch accent (such as a late peak after a H*).

2.12.4 Uncertainty about whether or not a syllable is pitch-accented

There will be cases of uncertainty relating to pitch accent that are not only about which type of pitch accent to use, but about whether to use a pitch accent label at all. The label *? is used for just such cases, and its use will be discussed below.

Levels of prominence:
As discussed in section 2.3, syllables in spoken language can be produced with a range of degrees of prominence: it is not the case, for example, that all syllables which do not bear a pitch accent are equally weak. Some syllables can be prominent compared to neighboring weaker syllables, but still less prominent than a syllable bearing a pitch accent. To make matters more complex, not all pitch-accented syllables are equally prominent either. For example, it may be the case that in an intonational phrase with two pitch accents, one pitch-accented syllable may be much perceptually stronger, and more salient than the other. The less-strong pitch-accented syllable is still pitch-accented, nevertheless.

There may be times, though, where it is not clear whether a syllable, which has some degree of prominence, is in fact pitch-accented. The example <august> shows an instance where the labeler was not certain whether a particular syllable was pitch-accented or not. In this example, shown in Figure 2.11.3, the labeler has used the *? label on the word *how* in the second intermediate phrase, *for how long* in that file. In this example, the labeler's use of the L* pitch accent symbol on the word *long* shows that she was certain that the word was pitch accented, and that the pitch patterns were best captured by the Low pitch accent. However, the labeler felt that the word *how* could also be pitch-accented. While not as prominent as *long*, the word seems to be fairly prominent. The pitch track also shows some ambiguity: there is a slight rise in f0 at the beginning of the word *how*. While this is potentially the result of a pitch-tracking error, the labeler may have been uncertain whether the rise in pitch, combined with the sense that the word *how* is fairly prominent, was evidence of a pitch accent on that syllable. The labeller addressed this uncertainty by using the *? label in the tones tier. Then, in the alt tier, the labeller put the type of pitch accent label that she would use had she felt confident that indeed the syllable was pitch-accented.
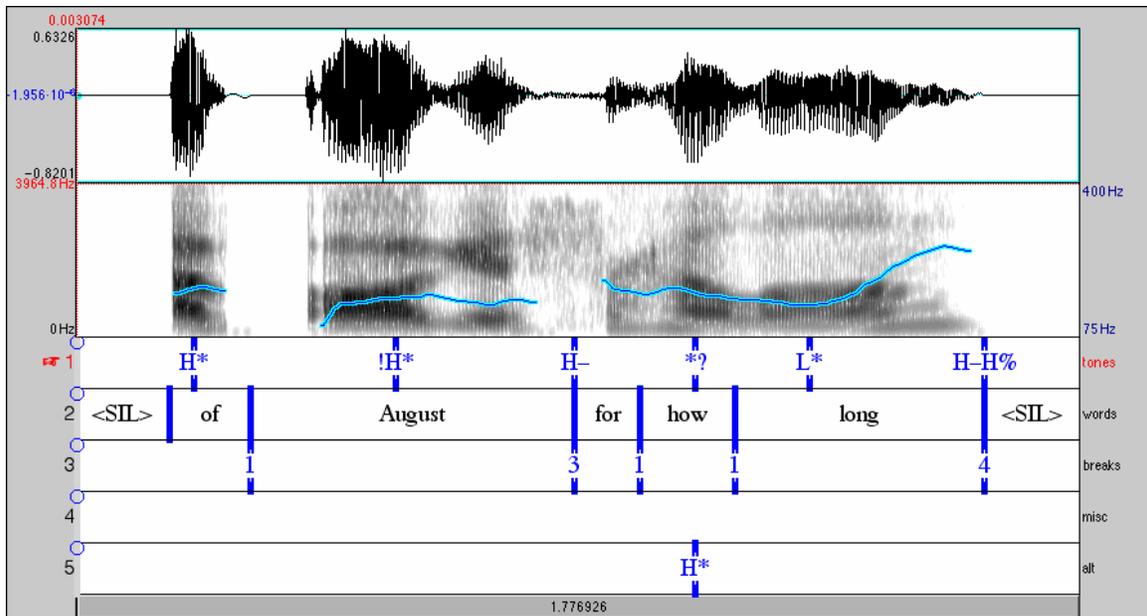
Figure 2.12.3 Using the alt tier and the *? symbol to indicate uncertainty about the choice of pitch accent label on *how*   <august>

Sources of uncertainty/ambiguity:

There are a number of fairly common phenomena that make it difficult to determine whether or not a certain syllable is pitch-accented. It is sometimes the case that a labeller will be uncertain whether a strong syllable between two (clearly) pitch-accented syllables also contains a pitch accent. This can happen when the syllable in question occurs during a stretch of high pitch between two H* pitch accents, or comparably during a stretch of low pitch between two L* pitch accents.

The example <marmalade7> shows an example of a *? label in a region between two L* pitch accents. In this example, there is a Low-tone pitch accent on the –an- of *Marianna*, and then another Low-tone pitch accent on the *mar-* of *marmalade*. The pitch lowers into the first of these two pitch accents, and stays low through the second. Notice how one of the syllables between these two pitch accents, the monosyllabic word *made*, is also fairly strong, and could potentially also have a Low (L*) pitch accent: the signal is ambiguous here. The labeller has captured this ambiguity, and resulting uncertainty, by using the *? label in the tones tier, and L* in the alt tier.
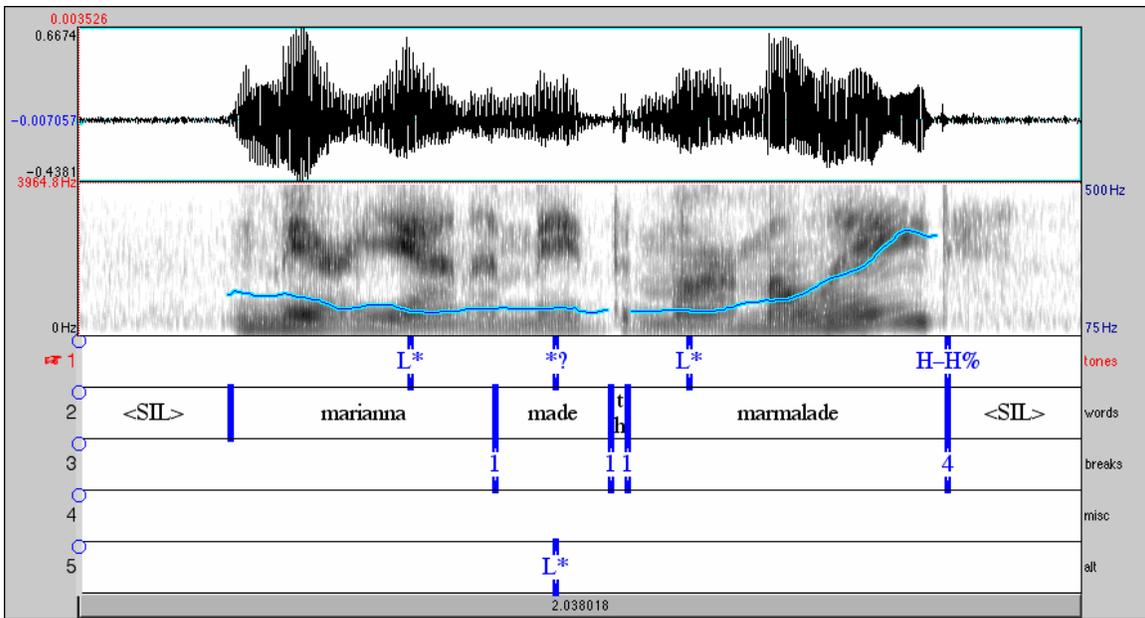
Figure 2.12.4   Using *? to mark uncertainty on *made*   <marmalade7>

The example <anna3>, shown in Figure 2.12.5, shows a parallel case to the one discussed above: in this case, a strong syllable produced in a stretch of high pitch between two syllables with High (H*) pitch accents could potentially also be pitch-accented with H*. Here, while the *An-* of *Anna* and the *Len-* of *Lenny* both are clearly pitch-accented (with H* pitch accents) the syllable *mar-* of *married* could also potentially be pitch accented; the label *? is used for this syllable in the tones tier, and H* in the alt tier.
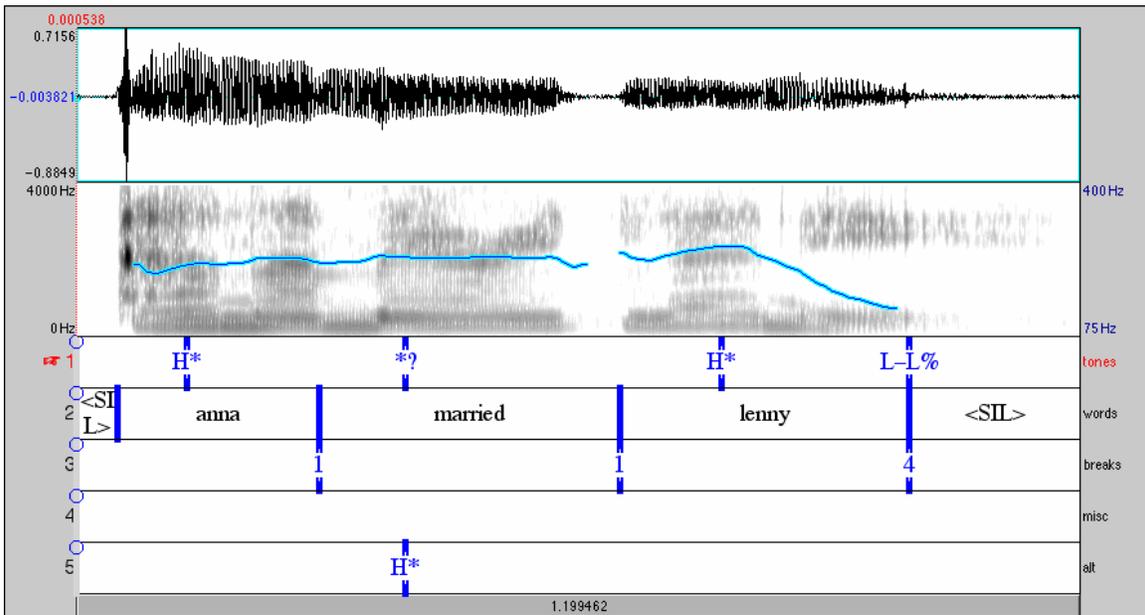

Figure 2.12.5   Using the alt tier and the ? symbol to indicate uncertainty about the choice of pitch accent label on *mar-* in *married*   <anna3>

Table 2.12.3 Other common contexts for *? (in addition to those discussed above):

• after one or more !H* pitch accents, such that the local pitch range has been greatly compressed
• when the speaker is generally using reduced pitch range
• when there is a phrase final full-vowel syllable, especially a monosyllabic word, where the duration cues to accent and boundary may be confounded.
• a phrase-initial full-vowel syllable, especially starting with a vowel (glottal stop)
• 2 consecutive full-vowel syllables: both seem strong, and it may be hard to tell which of the two is pitch accented—there may be times when there is pitch movement in the area of these two syllables that could be associated with an accent on one or the other or both.

2.12.5 Uncertainty regarding level of disjuncture

Labellers may also encounter times when they are uncertain about the appropriate break index. There are two ways which the labeller may signal uncertainty about level of disjuncture, or type of break index, which will be discussed in the following sections: 1) using the "minus" diacritic ("-") and 2) using the alternatives tier.

2.12.5.1 The "minus" diacritic:

The minus diacritic is used in conjunction with a break index number to indicate uncertainty about break index level where there are 2 clear alternatives, such as about whether break index 4 or break index 3 best captures the degree of disjuncture. Conventions for use of the "-" diacritic prescribe that in such cases, the higher break should be used in the breaks tier, with the "-" diacritic. So, uncertainty between break index level 4 and level 3 can be indicated with the break index of "4-". It is not necessary to use the alt tier for such cases, as the two alternative labels are understood to be 4 vs. 3. In such cases, the labeller must mark the phrase accent and boundary tone associated with the higher break level, the 4. In the example <sure>, shown in Figure 2.12.16, below, the labeller was uncertain whether the disjuncture between the words *sure* and *of* would best be captured by break index 4, indicating the end of a full intonational phrase, or whether break index 3, indicating an intermediate phrase boundary, would best fit. The lengthening in the word *sure* seems a bit short for a 4, but perhaps longer than might be expected for a 3. The labeller captured this uncertainty, and the ambiguity in the signal, by using the label 4- in the breaks tier. The labeller also put the corresponding phrase accent-boundary tone combination of L-L% in the tones tier, corresponding to a 4 break. The use of the 4- break index with L-L% in the tones tier indicates that the labeller considered that a break index 3, and a corresponding L- phrase accent, was a possible alternative.
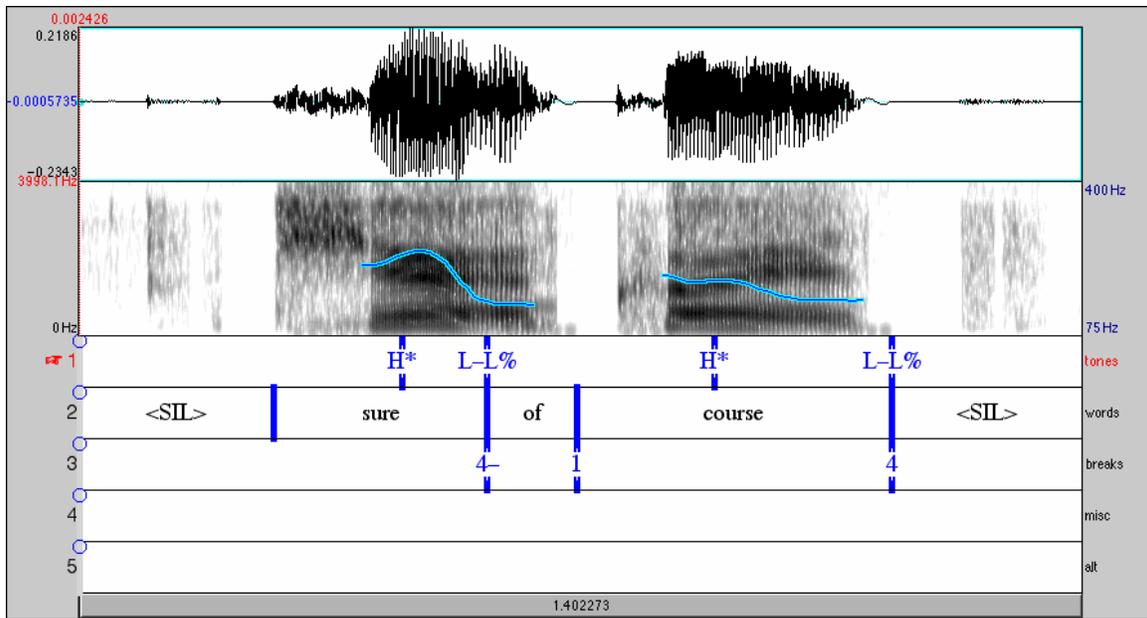
Figure 2.12.16 Using a 4- diacritic to indicate uncertainty about the level of break index on *sure* at the end of the first intonational phrase   <sure1> (version 1)

2.12.5.2 Using the alt tier for uncertainty about break index level

The use of the "minus" diacritic does not make it explicit which of the two alternatives the labeller preferred. For cases where the labeller wishes to indicate a strong preference for one break label over the other alternative, the labeller may make use of the alt tier. Figure 2.12.16, below, shows the same file, <sure>, with the alt tier used instead of the "minus" diacritic. In this case, the labeller shows a stronger preference for considering that the end of the word *sure* corresponds to the end of an intermediate phrase (rather than a full intonational phrase). The labeller has put break index 3 in the breaks tier, followed by the "?" diacritic, and the phrase accent L- in the tones tier, also followed by the "?" diacritic. Aligned to the time of both of these labels, the labeller puts a mark in the alt tier with both the alternative break index, and the corresponding phrase accent-boundary tone combination with the "&" ("ampersand") symbol between them: 4&L-L%.
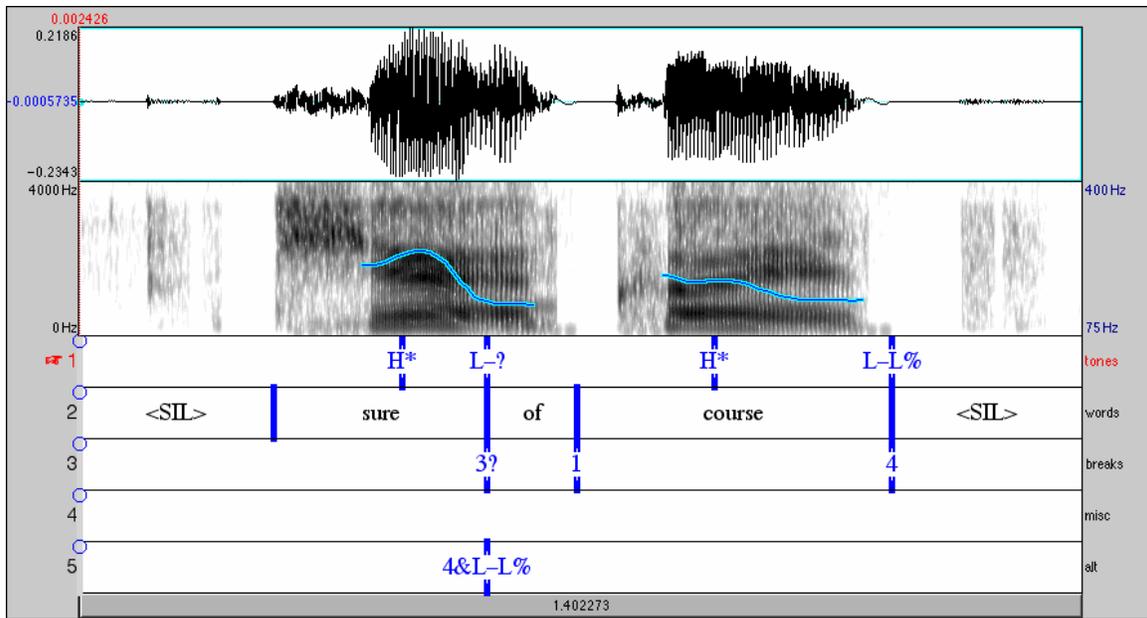
Figure 2.12.16 Using the alt tier and the ? symbol to indicate uncertainty about the choice of break index on *sure* at the end of the first intonational phrase

<div align="right"><sure2> (version 2)</div>

2.12.6 Uncertainty about phrase accent or boundary tone type:

There will be times when the labeller is uncertain what the appropriate phrase accent or boundary tone is for cases of intermediate and full intonational phrases. In such cases, labellers can again use the "?" diacritic with the tone markers in the tones tier, and list alternatives in the alt tier. In the example <diagonal>, the labeller indicated uncertainty about which boundary tone best would capture the tone patterns at the end of the final full intonational phrase of the file, realized on the word *mountain*. Here, the labeller was uncertain that the appropriate phrase accent was H-, but less certain about the appropriate boundary tone. While the pitch track seems to show a final rise, as would be characterized by H-H% following a H* pitch accent, the labeller perceived the pitch as being fairly flat, and not especially rising, as would be expected from the sequence H* H-L%. To capture that the pitch appeared to be rising more than expected, the labeller used the "?" diacritic after the phrase accent-boundary tone combination in the tones tier, and put H-H% in the alt tier. (Note that the labeller repeated the phrase accent part of the combination in the alt tier, rather than simply listing the boundary tone alone.)
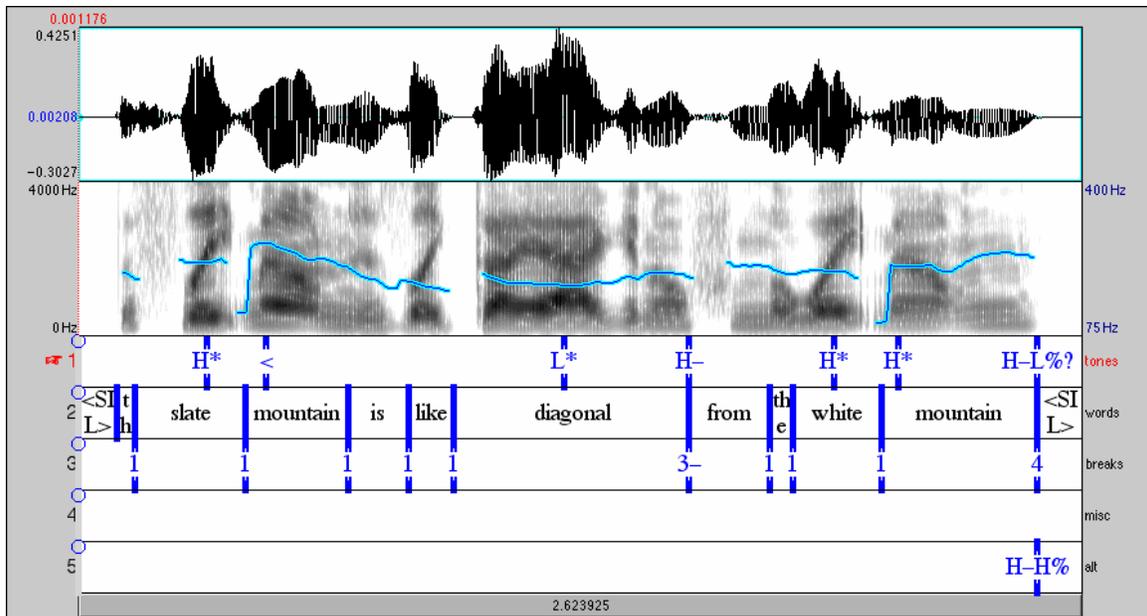
Figure 2.12.17 Using the alt tier and the ? symbol to indicate uncertainty about the choice of phrase accent and boundary tone at the end of the intonational phrase

<diagonal>

In cases where there is uncertainty about the phrase accent portion of a phrase accent-boundary tone combination, the labeller should put the "?" diacritic following the boundary tone in the tones tier, and list the alternative phrase accent type with boundary tone in the alt tier. (For example, if uncertain whether a phrase accent should be !H- or L- in a phrase accent-boundary tone combination, the labeller should put, for instance, !H-L%? in the tones tier and L-L% in the alt tier.)

2.12.6.1 X- and X%

ToBI has an additional two uncertainty markers, which can be used where the labeller has no "best guess" about which phrase accent or boundary tone to use: X-? and X%? respectively. These should be used only as a "last resort," such as in cases where the recording quality is very poor.

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 accent, low boundary tone
H-H%: high phrase accent, high boundary tone
L-H%: low phrase accent, high boundary tone
H-L%: high phrase accent, low boundary tone
!H-L%: downstepped high phrase accent, low boundary tone
H-: high phrase accent
L-: low phrase accent
!H-: downstepped high phrase accent

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

Optional labels:
<: late High Tonal peak

Uncertainty markers:

*?	uncertainty marker used on the tones tier to indicate that the labeller is not sure whether the specific syllalble is pitch-accented
X*?	uncertainty marker used on the tones tier to indicate that the syllable in question is definitely pitch-accented, but the labeller is unable to determine or decide which type of pitch accent
?	used in the tones or breaks tier to indicate that the Alternatives (alt) tier has been used for a particular label
-	used after a break index (on the breaks tier) number to indicate uncertainty between two levels of disjuncture (4-, 3-, 1-)
4-	uncertainty marker to indicate that the labeller is uncertain whether the break in question is a 3 or a 4. Both the phrase accent and boundary tone must be marked in these cases, and the break index should be considered a 3.
3-	uncertainty marker to indicate that the labeller is uncertain whether break should be 3 or a smaller index. The phrase accent must be marked.
1-	1 or 0
X-?	uncertainty marker to indicate that the labeller is unable to determine the type of phrase accent present at the end of an Intermediate Phrase (or in conjunction with a boundary tone, at the end of a full Intonational Phrase)
X%?	uncertainty marker to indicate that the labeller is unable to determine the type of boundary tone present at the end of a full Intonational Phrase
X-H%?	unable to determine whether L-H%, H-H% or !H-H%
X-L%?	unable to determine whether L-L%, H-L% or !H-L%

H-X%?	H-L% or H-H%
L-X%?	L-L% or L-H%
!H-X%?	!H-L% or !H-H%