

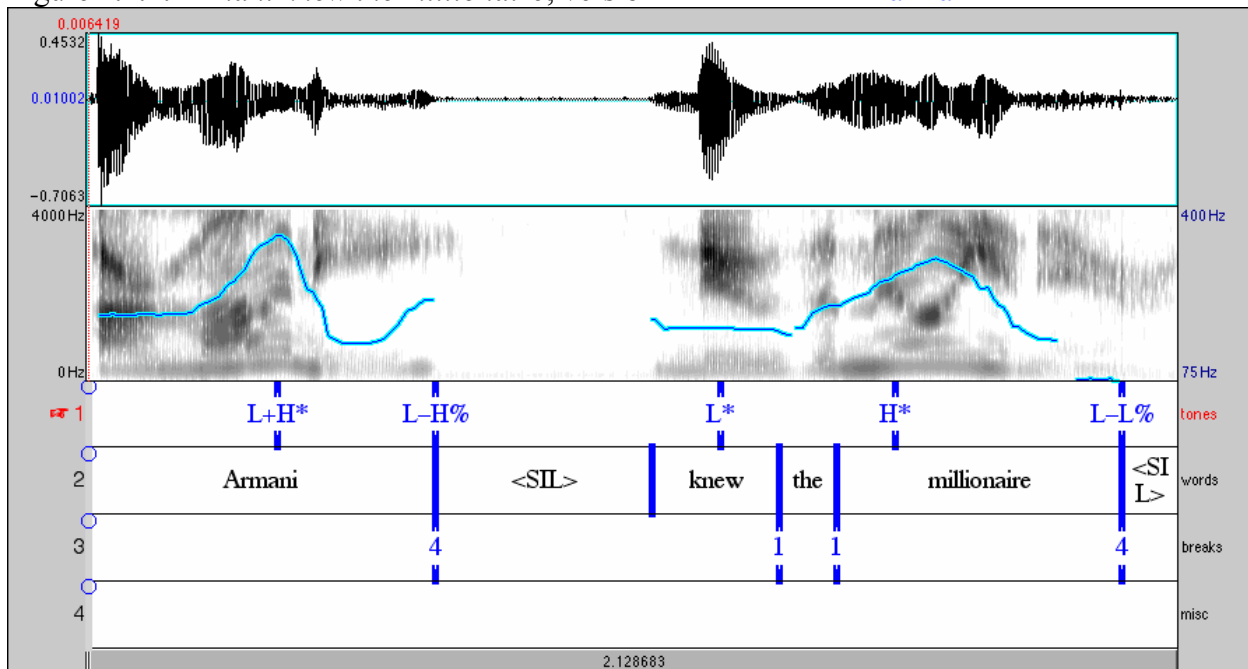
2.0 ToBI Tutorial

As we have seen, ToBI is a system of prosodic labelling that attempts to capture certain of the prosodic events in a spoken utterance. These prosodic events fall into two categories: 1) events or patterns that mark syllables as more salient or more prominent than neighboring syllables by reason of their intonation, and 2) events or patterns that mark phrasing or grouping within sequences of words.

For example, in the utterance shown in section 1 and repeated here <armani>, the words *Armani*, *knew* and *millionaire* are more prominent than *the*. Moreover, within the multi-syllable word *Armani*, the syllable *-man-* is more prominent than *Ar-* or *-ni*. (What is your intuition about the relative prominence of *mill-*, *-io-* and *-naire*?) The sentence is also produced as two intonational phrases. That is, the first word, *Armani*, is not grouped with the remaining words in the utterance, *knew the millionaire*.

Figure 2.1.1: *Armani knew the millionaire*, version 1

<armani>



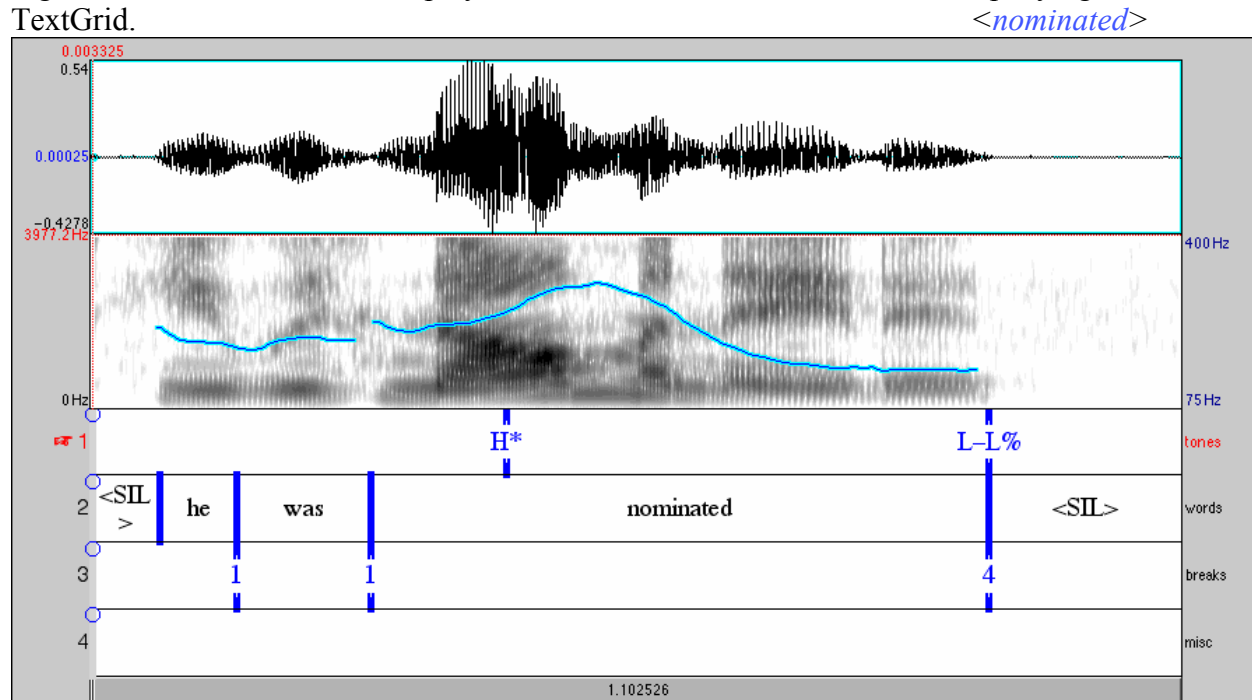
The acoustic correlates that mark these prominences and groupings are typically the f_0 pattern (related to the perceived pitch pattern) and the relative duration; the contributions of other cues such as amplitude (related to perceived loudness), voice quality, pausing, strength of articulation etc. are live topics of research and discussion. The Tone tier captures information about both phrasing and prominence, while the Break Index tier captures primarily phrasing and other grouping information. In the following sections of this introductory tutorial, elements from both of these tiers will be gradually introduced using examples that you can listen to and view labels for.

The first examples you will see are intentionally chosen as straightforward and context independent, to serve as clear illustrations of individual prosodic elements. As the tutorial progresses, more complexity will be introduced, showing the kinds of challenges that can arise in labelling when more elements are present to interact. Unfortunately, isolated examples heard out of context are sometimes missing substantial amounts of important information that can help the listener to parse (interpret) the cues in the acoustic signal in relationship to ToBI labels; examples of later sections will provide more context.

2.1 A first example, introducing the tone labels H*, L-L%, and the break labels 0, 1, and 4

One of the most basic prosodic patterns can be seen in the short utterance in <nominated>, shown in Figure 2, below. (To look closely at this example, open the files nominated.wav and nominated.TextGrid in Praat.)

Figure 2.1.2: nominated.wav displayed with nominated.TextGrid, the accompanying Praat TextGrid.

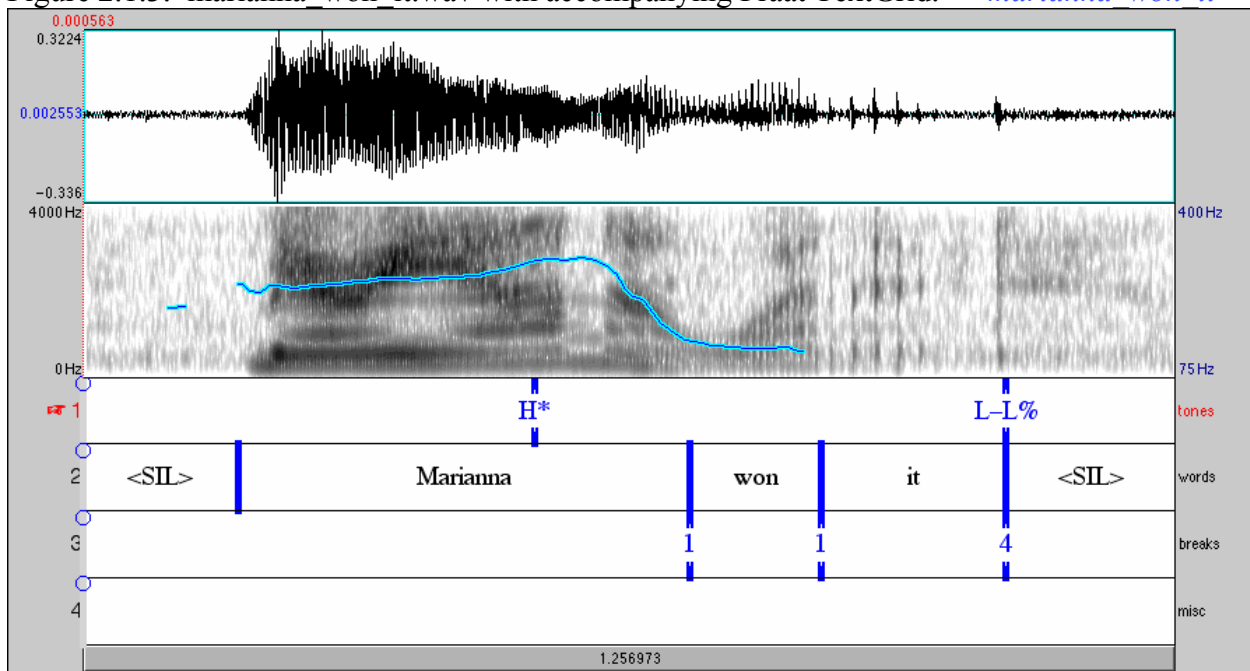


This example has 3 words, and is produced as a single intonational phrase with one pitch accent. Let's look at the transcription for the Tone tier first. As mentioned above, the Tone tier marks two types of events: those associated with accented syllables and those associated with phrasing. The most prominent syllable in this utterance is the first syllable (nominated) of nominated. The f0 contour shows a rise that peaks at the end of nom- and then falls. This is an example of a canonical H* or High pitch accent. In the ToBI transcription, the H stands for a High tonal target and the * indicates that the H tone is associated with an accented syllable. Perceptually, this H* syllable is more salient than other syllables around it. There are other types of pitch accents, including bi-tonal types, that make up the full ToBI inventory. We'll look at these other types in later sections.

The other intonational event that is labelled in the ToBI transcription of this utterance is the falling f0 at the end of the utterance. This low-pitched region is marked with L-L% where the L- indicates that there is a Low phrase accent that is followed by a Low boundary tone (indicated by the L%) on the final syllable. In later sections we'll look at phrase accents and boundary tones separately.

Like all intonation contours, this one can be produced on many different word strings. Another short example is shown in the file <marrianna_won_it>. Again, this contour is a single intonational phrase with one pitch accent. The pitch accent is an H* and the phrase accent/boundary tone combination is an L-L%.

Figure 2.1.3: marianna_won_it.wav with accompanying Praat TextGrid. <marrianna_won_it>

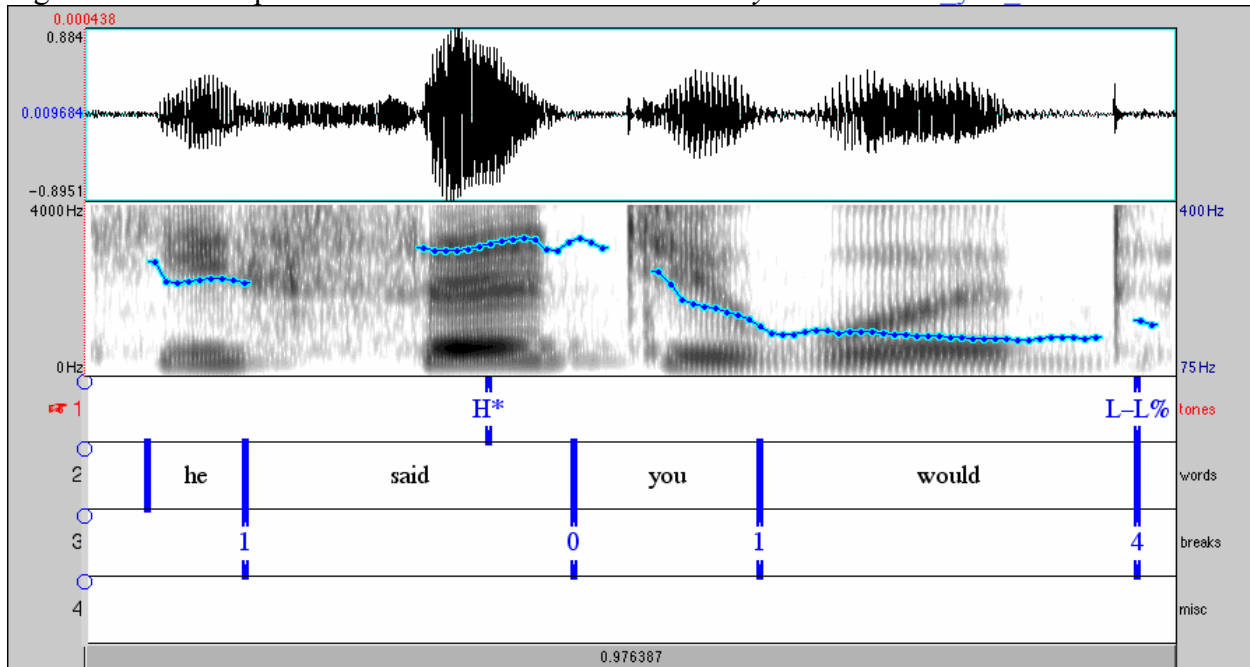


This L-L% sequence is often found at the end of spoken declarative sentences, especially final sentences in a turn or discourse. Later we will see phrase boundaries of a different type that have only a phrase accent. Moreover, just as pitch accents can be combined in different ways within an intonation contour, there are other combinations of high and low tones that make up the full inventory of phrase-boundary-related tonal markers.

Before introducing other kinds of pitch accents, phrase accents and boundary tones, however, let's turn our attention to labels in the Break Index tier. A Break Index marks the level of disjuncture between two words. Since intonation is one of the acoustic cues to disjuncture, there is some redundancy between the Tone and Break Index tiers. For example, an L-L% boundary tone signals the maximal level of disjuncture in this system (i.e. at the end of an Intonational Phrase) and would be marked with a 4. Other boundaries are marked with smaller break indices. For example, typical word boundaries in a fluent sequence of words within a phrase are marked

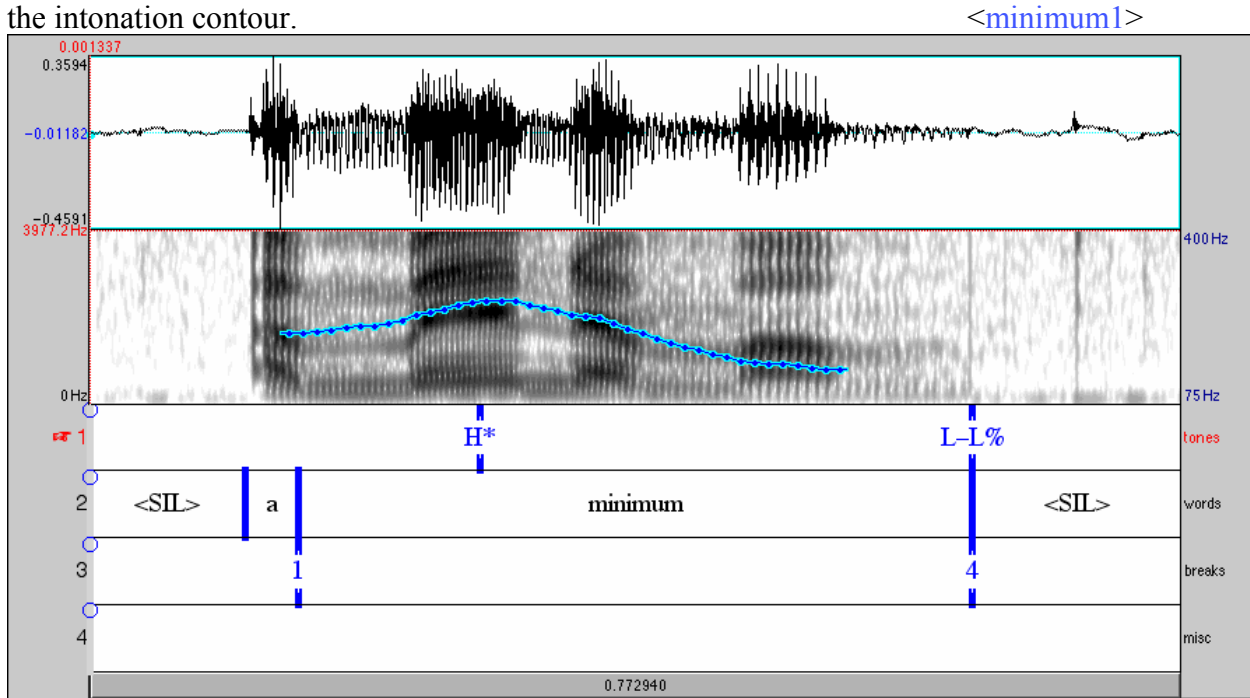
with a 1. In this example, all break indices are 1's except for the final 4 break index. There are levels of intermediate phrasing within intonational phrases that will be marked with break indices 2 or 3, as we will show later. The 0 break index is reserved for the case when two words are produced so that the boundary between them is indeterminate. For example, the word sequence *did you* is frequently produced as *didja*, where the final /d/ of *did* and the initial /y/ of *you* together form a /j/-like sound. (In example <said_you_would> “He said you would”, listen for the merging of the final /d/ of *said* and the initial /y/ of *you*.)

Figure 2.1.3: Example of Break Index 0 between *said* and *you*. <said_you_would>



The following examples show more instances of the same H* L-L% contour on utterances of varying lengths.

Figure 2.1.4: example <minimum1>, a 2-word utterance with H* L-L%. Here, the initial low tone might suggest an alternate pitch accent (L+H*) that will be discussed later. A newly adopted tier, the Alternatives tier, allows labellers to capture these reasonable alternate parses of the intonation contour.



To summarize the inventory that we've seen so far:

Tones:

H* high pitch accent

L-L% low phrase accent, low boundary tone

Break indices:

0: word boundary apparently erased

1: typical between-word disjuncture within a phrase

4: end of an intonational phrase