A phonological model of English intonation – Groundwork for English-ToBI  
(e.g., Pierrehumbert, 1980; Beckman & Pierrehumbert, 1986)

“Intonation has a phonological organization” -- based on autosegmental and metrical phonology.

Two aims of Pierrehumbert (1980):

1. To develop an abstract representation for English intonation.
2. To investigate the rules which map these phonological representations into phonetic representations.
   - Pierrehumbert wanted to identify the properties of the surface representation that can be explained by rules applied during the derivation instead of being marked in the underlying form—i.e., what is the simplest possible underlying representation that can be assumed?

A. Three components to the phonology of intonation:

1. A grammar of allowable phrasal tunes (=syntax of tones): this grammar generates sequences of L and H tones, as shown in Pierrehumbert’s: finite state grammar:

(Figure, revised from Pierrehumbert 1980, reflecting the changes in Beckman & Pierrehumbert 1986 model; Tones in parenthesis exist only in Pierrehumbert’s (1980) model. A long line ending in a filled arrowhead shows that pitch accents followed by phrase accent H-/L- can be repeated.)
Phrase accents (-) and boundary tones (%) signal a prosodic structure.

The figure above (Beckman and Pierrehumbert’s 1986 model) shows:

- The beginning of an IP is optionally marked by %H boundary tone.
- The end of an IP is always marked by a boundary tone (H% or L%).
- The end of an ip is always marked by a phrase accent (H- or L-).
- An IP always has at least one ip.
- An ip can have at least one pitch accent.
- There are six pitch accent types.

- H- or L- phrase accent marks the end of an intermediate phrase (ip) and H% or L% boundary tone marks the end of an Intonation Phrase (IP) (the prosodic unit, ip, is proposed first in Beckman & Pierrehumbert 1986, not Pierrehumbert’s original 1980 model). An IP contains one or more ip’s, and an ip contains one or more words with pitch accent.

- If a phrase has only one syllable, the syllable is phonologically represented with three different types of tones (pitch accent (*), phrase accent (-), and boundary tone (%)). See the hierarchical structure below (right side).

Hierarchical prosodic structure for the study of English intonation

\[
\begin{align*}
\text{IP} & = \text{Intonation Phrase} \\
\text{ip} & = \text{intermediate phrase} \\
\text{W} & = \text{word,} \\
\sigma & = \text{stressed syllable} \\
\ldots & = \text{unstressed syllable(s)} \\
\end{align*}
\]

\[
\begin{align*}
T & = \text{tones (H or L or combination of H & L)} \\
T* & = \text{pitch accent} \\
T\% & = \text{boundary tone (linked to IP)} \\
T– & = \text{phrase accent (linked to ip)} \\
\end{align*}
\]
2. A representation of the text using a metrical grid

Prominence relations as a metrical representation

- In many languages, (e.g., English), there are prominence relations between syllables within a word (lexical stress), and prominence relations between words within a phrase (phrasal, or “post-lexical”, stress/accent). These relations can be described as *Metrical Phonology*.

- Even though every word in English has stress at the lexical level, not every word is prominent at a phrasal/post-lexical level. Which word receives prominence is not fully predictable, but it’s not completely unpredictable either (contra Bolinger 1972).

- Prominence is known to be partially determined by the semantic and pragmatic meaning of the utterance, speaker intention, as well as by rhythmic constraints, lexical frequency, and parts of speech. (Calhoun (2006).

- Prominence at the level of sentence-level prosody/intonation is treated as a level of stress, higher than the content word stress level. Stress is a relational, not absolute, term; higher-level prominence structure works the same way.

- Prominence relations can be represented using the metrical grid of Liberman (1975) and Liberman & Prince (1977) (see also Beckman 1986):

  **Ex1:**
  
  \[
  \begin{array}{cccccccc}
  & x & & & & & & \\
  x & x & & & & & & \\
  x & x & x & x & x & & & \\
  x & x & x & x & x & x & x & x \\
  \end{array}
  \]

  Legumes are a GOOD source of vitamins

  **Ex2:**
  
  \[
  \begin{array}{cccccccc}
  & x & & & & & & \\
  x & x & & & & & & \\
  x & x & x & x & x & x & & & \\
  \end{array}
  \]

  That’s a remarkably clever suggestion

- A pitch-accented syllable (a lexically-stressed syllable aligned with a tone) is more prominent than a lexically stressed but non-pitch-accented syllable. Thus, pitch accent is a higher degree of prominence.
• The last pitch accent in an ip, called the *nuclear pitch accent*, is more prominent than earlier pitch accents, called *prenuclear pitch accent*, suggesting a hierarchical relation of prominence.

• Note that this implies that prominence structure is closely tied to phrasing structure; *Manitowac* is prenuclear in Ex3, but nuclear in Ex4 (*library* is nuclear in both):

Ex3:  
```
[[(Does Manitowac have a library?)_ip]_IP
          x                  Nuclear pitch accent
          x                  Pitch accent
          x                  Content word stress
          x  x  x  x  x  x  x  x  x  Syllable
Does Manitowac have a library
```

Ex4:  
```
[[(Does Manitowac]_ip [have a library?]_ip]_IP
          x                  Nuclear pitch accent
          x                  Pitch accent
          x                  Content word stress
          x  x  x  x  x  x  x  Syllable
Does Manitowac have a library
```

3. Rules for lining up the tune with the text:

• A complete phonological representation for intonation is a *metrical representation of the text with tones aligned according to rules*.

Ex.  
```
x                  x                  Nuclear pitch accent
x                  x                  Pitch accent
x                  x                  Content word stress
x  x  x  x  x  x  x  x  x  Syllable
My brother returned the book from the library.
          |                  |                  |
          H*              L-                H*  H*  L-  L%
(Intermediate phrase) (Intermediate phrase)
(--------------- Intonational Phrase ---------------)
Revision from Pierrehumbert (1980) to Beckman & Pierrehumbert (1986) to ToBI (1994)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>7 Pitch Accents</td>
<td>6 Pitch Accents</td>
<td>5 Pitch Accents</td>
</tr>
<tr>
<td>H*</td>
<td>H*</td>
<td>H* (and !H*)</td>
</tr>
<tr>
<td>L*</td>
<td>L*</td>
<td>L*</td>
</tr>
<tr>
<td>H*+L*</td>
<td>H+L*</td>
<td>H+H*</td>
</tr>
<tr>
<td>H*+L-</td>
<td>H*+L</td>
<td>L*+H (and (L*+H))</td>
</tr>
<tr>
<td>L*+H-</td>
<td>L*+H</td>
<td>L+H* (and L+H*)</td>
</tr>
<tr>
<td>L^+H*</td>
<td>L+H*</td>
<td></td>
</tr>
<tr>
<td>H^+H-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phrase accent (H-, L-), but no Intermediate phrase level</td>
<td>Intermediate phrase tone (H-, L-)</td>
<td>Intermediate phrase tone (H-, L-, !H-)</td>
</tr>
<tr>
<td>Boundary tone (X%)</td>
<td>Intonational phrase tone (X%)</td>
<td>Inton. phrase tone (%X, X%)</td>
</tr>
<tr>
<td>The tag in Tag Question has no pitch accent (X- X%)</td>
<td>The tag in Tag Questions have a pitch accent: X* X- X%</td>
<td>(because every ip should have at least one pitch accent.)</td>
</tr>
</tbody>
</table>

Three independent parameters of intonation:

i. Pitch Accent Placement (i.e. which word(s) get pitch accented?)

ii. Pitch Accent Type (i.e. what tone is aligned with the word’s stressed syllable?)

iii. Phrasing (i.e. which words are grouped together—and their edges marked by tones?)

On pitch-accenting (and this section leads into the Beckman & Edwards paper):

- In general, a content word (optionally—depending in part on sentence meaning) gets one pitch accent. A longer word, however, can have two or more pitch accents; these prenuclear accents align with secondary stressed syllables:

Ex:  
   a. CAL-i-FORN-ia
   b. BA-lle-RI-na
   c. MEL-a-NO-ma

⇒ When the secondary stress syllable precedes the primary stressed syllable, it is the secondary stress one that gets a prenuclear accent; the primary stressed syllable bears the nuclear pitch accent.
Note that if the primary stress precedes the (putative) secondary stress, that secondary stress cannot carry a pitch accent in English (instead there is just one pitch accent—the nuclear pitch accent on the primary stressed syllable):

Ex:  
   a. EL-e-ve-tor
   b. OP-er-a-tor
   c. CAT-er-pill-ar

- It has been suggested that what has traditionally been described as cases of “stress shift” are better analyzed as “prenuclear pitch accent” shifting. (See Beckman and Edwards’ *Chinese men* example). Classic cases include those subsumed under the “rhythm rule”, which moves prominence rightward in *thirteen* when a stressed syllable follows in, for example, *men*:

  **The Rhythm Rule (e.g., Hayes 1995):**

  a. thir-TEEN

  *but:*

  b. THIR-teen MEN

- However, there are also cases where it seems necessary to assume the “promotion” of prominence at the lexical stress level too (i.e., a non-full vowel gets promoted to full vowel, which then gets promoted to a pitch-accented vowel). Such cases are generally the result of contrastive focus (either linguistic focus or paralinguistic emphasis):

  Ex:  
  ...I didn’t say he needs to DEport it, I said he needs to EXport it.

  ➤ Contrast/emphasis changes the syllable associated with the *-tone (assuming that the verbs (not nouns) “deport” and “export” have stress on the second syllable for most speakers).