Tone & Intonation
Tone

- Linguistic uses of pitch (f0) gestures
- Two contrasting tones (or pitch gestures): H, L
- Complex tonal patterns can be decomposed into sequences of H and L.
- Tone gestures and sequences of tone gestures can be coordinated with (coupled to):
  - syllables (“lexical tone”) — distinguish lexical items
  - words (“pitch accent”) — distinguish lexical items
  - phrases (“intonation”) — distinguish syntactic and pragmatic structures.
Assignment to Syllables: Register Tones

- can be used to contrast lexical items
- used in a majority of the world’s languages

Shona (Zimbabwe)

[kùtʃérá] ‘to draw water’
[kùtʃèrə] ‘to dig’

Bini (Nigeria)

[i mà] ‘I show’
[i mà] ‘I am showing’
[i má] ‘I showed’

Yoruba (Nigeria)

Hombert, 1976

up to 5 contrastive tone levels: how do they result from H, L??
H and L tones in Embosi  
(Bantu Language in Congo)

(6) [ikósibílámbrinéébílalapóimisóβámina]  
(b)i-kós bílámbrí neebí la (m)a-póa (b)í-misáá  
cl8-manioc cl8.REL.cook.REC Gneebii at cl6-yesterday cl8-already.is  
ó-βámina  
cl15-toughen  
‘The manioc that cooked Gneebii yesterday is already toughened.’

Figure 3: F0 curve of [ikósibílámbrinéébílalapóimisóβámina] “The manioc that Gneebii cooked yesterday is already toughened” (Speaker MEA)
Contours

- Multiple pitch gesture per syllable: contour tone
- Pitch gestures may not have their own syllable to be coupled to.
- Or, their coupled vowels may be elided in running speech when adjacent to another vowel.
- They can then be coordinated with a preceding or following syllable creating a tonal sequence.

Etsako (Nigeria)

\[
\begin{align*}
\text{he} & \quad \text{PAST} & \quad \text{buy} & \quad \text{cup} & \quad \rightarrow & \quad [\ddot{o}d\dot{u}k\ddot{e}k\ddot{w}\ddot{i}] & \quad \text{‘he bought a cup’} \\
\text{he} & \quad \text{PAST} & \quad \text{buy} & \quad \text{chair} & \quad \rightarrow & \quad [\ddot{o}d\dot{u}t\ddot{e}k\ddot{w}\ddot{i}] & \quad \text{‘he bought a chair’}
\end{align*}
\]
Shekgalagari (Bantu, Botswana)
(Hyman & Monaka, 2011)
Assignment of Tone Sequences to Syllables: *Contour Tones*

- Tonal sequence can contrast lexical items

**Standard Chinese**

<table>
<thead>
<tr>
<th>Tonal Sequence</th>
<th>Tone</th>
<th>Syllable Symbol</th>
<th>Pitch (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High level</td>
<td>L</td>
<td>ma 55</td>
<td>350</td>
</tr>
<tr>
<td>High rising</td>
<td>H</td>
<td>ma 35</td>
<td>250</td>
</tr>
<tr>
<td>Low falling-rising</td>
<td>L</td>
<td>ma 214</td>
<td>150</td>
</tr>
<tr>
<td>High falling</td>
<td>L</td>
<td>ma 51</td>
<td>50</td>
</tr>
</tbody>
</table>


**Gestural Scores of Mandarin Tones**

- Tone 1: H
- Tone 2: L, H
- Tone 3: L
- Tone 4: H, L

Gao (2006)

![Pitch contours output by Task Dynamic Model (Nam et al., 2005)](image_url)
Mandarin Coupling Graphs

- Gao (2008) measured
  - onset of C,V gestures (using EMMA)
  - onset of pitch gesture (from turning point in f0)

- The relative timing of C, V, and Tones shows the pattern predicted by the competitive coupling of onset Cs.

![Diagram of Mandarin Coupling Graphs]

- Thus, the Mandarin tones are behaving like onset Cs, at least with respect to their timing.
Coordination of Mandarin Tones with C,V

Tone 1

Tone 3

Tone 4

Tone 4
Assignment of Tone Sequences to Words: *Pitch Accents*

- Words can contrast in terms of pitch gestures assigned to them, but not every syllable can independently bear pitch gestures.

  **Croatian**

  - Words contrast in accent type: 2 sequences of pitch gestures.
    - Early rise \((L+H^* L-L\%\) /mlada/ 'the bride'
    - Late rise \((L^*+H L-L\%\) /mara/ proper name ('Mara')
  - Patterns are "strung out" over words of 2- or 3-syllable words.
  - In a syllable tone language, there could be 8 contrastive tonal patterns on 3-syllable words. In Serbo-Croatian, there are only
Tone Gesture Control

- Gestures are defined by tasks and the articulators that produce them.

  - **Tasks:** F0 value (H or L)
  
  - **Articulators:**

    - Increase in longitudinal tension
    - Produced by increasing angle between cricoid and thyroid cartilages
    - Action of *crico-thyroid* muscles

    - Decrease in vertical tension produced by lowering entire larynx.
    - Action of the *strap* muscles: sterno-hyoid, thyrohyoid
EMG of Thai tones (Erickson, 2011)
Accent Patterns in Kinki Dialect of Japanese

(type A) - (type B) - (type C) - (type D)

(type A) - (type B) - (type C) - (type D)
Cantonese Tones

- Four pitch levels
- Nissenbaum (2010) hypothesized they are produced with a 2x2 combination of
  - Larynx Height (SH) (“register”)
  - CT Stretching

Cine-MRI evidence: extreme tones

Male speaker age 20

Upper and Lower extreme tones

UPPER register, HIGH tone

Hi Larynx
Stretched

LOWER register, LOW tone

Lo Larynx
Short Folds

Subjects:
Nine subjects participated (6 male, 3 female).
Subjects were all native speakers of Cantonese between the ages of 19 and 54.

Preliminary results from the MRI study

Male speaker age 20

Upper and Lower extreme tones

Mid tones

UPPER register, L-tone ("high mid" tone 3)

LOWER register, H-tone ("low mid" tone 6)
• In running speech, the f0 of the two middle tones are not distinct, but they are produced with distinct gesture combinations.

Mid tones

/उ3/  UPPER mid-tone
vocal fold length = 17.7 mm
posterior vertical dist. from top = 127.7 mm

/उ6/  LOWER mid-tone
vocal fold length = 19.7 mm
posterior vertical dist. from top = 132 mm

• Difference in vocal fold length between tones 3 and 6 for this subject is 2 mm (i.e. vocal folds are 11% longer at onset of tone 6 than at onset of tone 3)

• Difference in vertical position is 4.3 mm (i.e. larynx lowers by nearly 1/2 cm for tone 6)

Conclusions

Two production mechanisms are known to be available in principle for adjusting pitch independently. Characterizing tone and register features in terms of these mechanisms does better justice to the facts of Cantonese speech. The articulatory model is supported by high-speed MRI of Cantonese speakers. It has proven fruitful in many cases to characterize aspects of phonological knowledge in terms of the articulatory dimension of speech. When it comes to intonation (and other aspects of sound structure under the control of the larynx), linguists have tended overwhelmingly to characterize theoretical primitives in terms of acoustic dimensions (f0, VOT, etc.). These results suggest that there is something to be gained by taking the articulatory dimension of tone to be theoretically primary. A non-invasive technique for investigating aspects of speech production that have resisted empirical study, potentially making a broad range of questions available for novel methods of inquiry.
Intrinsic f0

- High vowels have higher F0 than low vowels (31 languages in Whalen & Leavitt, 1995).

- Effect has been considered universal and a “mechanical” effect of vowel production.
  - **Mechanism?** Low vowels pull the hyoid bone down, which also lowers the larynx.

- Counter-evidence to “mechanical”:
  - Effect is neutralized in Mambila, a 4-tone African language (Connell, 2002).
  - Regional variation in size of the effect in English (Jacewicz & Fox, 2015).
Voicing and F0

- Near-universal effect of voicing in stops and fricatives on f0
  - F0 is high and falling after voiceless stops and fricatives.
  - F0 is low and rising after voiced stops and fricatives.

- Mechanism unclear
  - LH for voiced stops causes lower F0 (but implosives may not have effect)
  - CT for voiceless stops contributes to devoicing.
  - Inconsistent patterns for nasal and pre-nasalized stops, murmured stops (Chibelli, 2015).

- Tonogenesis (Hyman, 1976)
  - Tone contrasts can develop and replace voicing contrast
Assignment of Tones & Sequences to Phrases: *Intonation*

- **Functions**
- **Sentential meaning (speech-act)**
  - assertion, question, command
- **Information structure, e.g., focus**
  - What about Manny? Who came with him?
    - Anna came with Manny.
  - What about Anna? Who did she come with?
    - Anna came with Manny.
- **Syntactic parsing**
  - "light-house keeping" vs. "light house-keeping"
  - "1 + (3 *2)" ... = 7
  - "(1 + 3) * 2" ... = 8
ToBI system of Intonation description

• Pitch accents (tones & sequences):
  Every phase has at least one.

• Final one is called the "nuclear" accent.

• (Partial) Inventory of pitch accents:
  • starred tone is coordinated with stressed vowel
  • !H means lowered H

• Phrase accents
  Can be added to the "nuclear" accent.

• Boundary tones
  • Final rising or falling pitch
Example of ToBI

- **H* L-L%**
  - Statement

- **L* L-H%**
  - Address

- **L* H-H%**
  - Question

- **L+H* L-H%**
  - Puzzlement

- **L+H* L-L%**
  - Reprimand
Hypothesis (Liberman & Pierrehumbert, 1984)

Prominence relation between (B and A) is invariant (across positions and and pitch range) though obscured by downstep and final lowering.

Recording of these sentences under 10 degrees of “overall emphasis” (to vary pitch range).
Downstep

- Recording of lists: 2-5 items
- Three levels of pitch range

Hypotheses

Downstep is can be modeled by an exponential decay function (first-order dynamical system with “goal” at bottom of pitch range).

Downstep is invariant across pitch range changes (when a reference level is incorporated that can vary across pitch ranges)
Results

- Results support exponential decay
- But final position is particularly low

Figure 21
Pitch range 3, length 5 data for subject DWS. The solid line is a decaying exponential fit to the first four points of the five-item list. The arrow indicates how far the fifth data point falls below the value predicted by this exponential.
Intonation in Lexical Tone Languages

- Theoretical problem: how are functions of intonation expressed in the presence of lexical tones
  - Morphemes
  - Intonational tone gestures that blend with lexical tone gestures
  - Pitch range
  - downstep or declination
4.1. Corpus design

The test sentences used in this experiment fell into three sets: all-H sentences, all-L sentences, and mixed-HL sentences similar to those used in Experiment 1. These sentences had a syntactic structure similar to those used in the earlier experiment. There were three length conditions in each set: (a) long, (b) medium, and (c) short. The test sequences were placed at the beginning of carrier sentences to absorb any final lowering effects. The sentences used in this experiment are listed in Appendix B. All sentences were read by TJ and FA and recorded and analyzed as described earlier.

4.2. Results

Fig. 5 overlays the graphs of mixed-HL sentences with those of the all-H and all-L sentences (tones of the sentence-final carrier frames are omitted here and in later figures). Graphs for TJ are shown on the left and those for FA on the right. Sentences are presented from shortest (a) to longest (f).

Downstep

- Successive H tones (separated by L) are lowered
- Can be modeled with an attractor to a low level at phrase end, with resets if the f0 gets too low

Yoruba

Declination in Mambila (Connell, 2016)

- 4 level tones
- f0 lowers over time, even for same-tone sequences
- T1 (H) lowers least
- T4 (L) lowers most

Finally, the results of T-tests comparing the two slopes (11) and the mean F0 of the two portions of the sentences (12) are both significant:

(11) Difference of slopes = 18.06 StdError = 4.7316
T-value = 3.8169, p-value = 0.0001

(12) Standard Two-Sample t-Test
t = 4.6377 , df = 28, p-value = 0.0001
Bemba (Zambia)

- Both Declaratives and Yes/no Questions have final L%.
- Declaratives exhibit downstep over the entire utterance.
- Yes/No Questions suspend downstep.

**Figure 13:** Pitch tracks of a declarative sentence (black) and the corresponding polar question (grey) (16)

(16) bànamáâyó bà-ká-’péēl-à úmu’kááte
2woman 2SM-FUT3-give-FV 3bread
‘The women will give the bread?’
Declaratives have final lowering over the penultimate and final syllables (L%), that is suspended in YNQ.

mulimi ubanelava kana bitari

Does the farmer want to leave?