1. syllable

- no agreed upon phonetic definition of syllable
- concept is inferred rather than perceived (Steriade 2000)
- syllabic writing systems are ubiquitous vs. alphabetic
- suggests that parsing speech into syllables is easier than into phonemes (e.g. Japanese hiragana)
- many verse structures built on number of syllables (French alexandrine: 12 syllables divided by a pause (caesura) between 6th and 7th; Berber 12 syllable lines with heavy syllable in 3rd, 7th, and 12th position (Dell & Elmediou 2002)

2. internal structure: onset-rhyme (nucleus coda)

- structure preservation in spoonerisms: our dear queen -> our queer dean break at onset.rhyme
- cross-linguistically rhyme is domain for poetic rhyme; rhyme books crucial to reconstruction of Chinese;
- open syllable ends in a vowel and closed syllable ends in a consonant
- phonetic definition in terms of chest pulses (Stetson) not confirmed
- peaks of sonority: sonority is roughly the acoustic intensity of a sound with loudness being a primary auditory correlate.
- sonority hierarchy: vowels (low,mid,high) > glides > r > liquids > nasals > fricatives > stops (voiced > voiceless) see CIP 10.1
- correlates with degree of opening of vocal tract
- no salient peaks lead to less robust judgments of syllable number: prism, file
- consonantal onset and coda are frequent sites of phonotactic restrictions (syllabic licensing): in English the velar nasal [ŋ] is barred from the onset while in nonrhotic dialects [r] is banned from the coda

[3] syllable parsing

- given VCV other things being equal, V.CV parse is preferrred
- Fujimura's (1978) experiment starts with cluster with different points of articulation: apta, atpa, then shorten closure duration to that of a single stop—but formant transition cues conflict: CV transitions predominate over VC
- judgments of syllabic division are inferred rather than based on a direct phonetic percept.
- judgements based on possible word initial and word final clusters
- languages with no clusters at beginning and end of word will divide cluster: CVC.CVC (e.g. Korean). cf. English, Romance where V.CR (ál.ge.bra) but VR.CV (al.bér.ta)
• #bra is a possible word initial cluster while #rta is not
• syllabification parse here may be analogous to treating each syllable as if it were a word
• but if we get a consonant that is not permitted initially then syllabification goes against the Onset preference. Cf. Korean no velar nasal initially. VŋV is parsed Vŋ.V
  tam        can        k*ŋ
  tam-i      can-i      k*ŋ-i
  'wall'     'cup'     'pheasant'

• cf. English lamént vs. lémon: [lɛ] not possible word-finally; thus lɛ.mon is not a good parse and [m] is perceived as ambisyllabic (straddles syllable boundary and belongs to both syllables)

[4]. Stress: will discuss later: phonetic correlates include greater muscular energy (subglottal pressure) with increased pitch and duration

[5]. Length

• long vs. short vowels: Latin, Czech, German; ratio of duration contrast is 1.3 : 1 up to 2: 1
• long (geminate) consonant: Italian 2 to 3 times longer
• both long vowels and consonants: Hungarian, Arabic, Japanese, Finnish
• inherent duration of vowels: low vowel [a] is 20-25 ms. longer than high vowel [i]
• in a CVC syllable more time is required for tongue to move from consonantal constriction to open low target in CVC syllable
• global effects:
  ➢ the number of segments in a syllable and the number of syllables in a word can decrease the length of the individual segments
  ➢ phrase-final lengthening
  ➢ tempo: faster tempo compresses longer segments (vowels) more

• Many languages avoid long vowel in closed syllable
  ➢ Italian lengthens stressed vowels of nonfinal syllable, but not when consonant in coda
    fate   'do'   fat.te   'done'
    177 ms.  126 ms.  D'Imperio & Rosenthal 1999

  ➢ Egyptian Arabic non-final closed-syllable shortening
    baab   'door'   beet   'house'
    bab-na  'our door'   bit-na  'our house'

  ➢ English: keep, kept
Cf. Hungarian, which retains contrast

<table>
<thead>
<tr>
<th></th>
<th>V</th>
<th>C</th>
<th>duration in ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>hal</td>
<td>151</td>
<td>182</td>
<td></td>
</tr>
<tr>
<td>hall</td>
<td>177</td>
<td>272</td>
<td></td>
</tr>
<tr>
<td>a:l</td>
<td>285</td>
<td>172</td>
<td></td>
</tr>
<tr>
<td>ne:z</td>
<td>287</td>
<td>198</td>
<td></td>
</tr>
<tr>
<td>ne:zz</td>
<td>263</td>
<td>288</td>
<td></td>
</tr>
</tbody>
</table>

[6] **tone:** pitch $F_0$
- gender: average of 130 Hz for adult males and 220 for females
- intrinsic $F_0$: high vowels greater by c. 15 Hz (JND for speech is 10 Hz)
- not clear why; speaker control to enhance height contrast?
- after voiceless obstruent $F_0$ is 5-10 Hz greater (basis of tonogenesis); again apparently under speaker control to enhance voicing contrast

[7] **tonogenesis**
- in many Asian languages tones derive historically from the phonemicization of redundant $F_0$ differences that accompany voiced vs. voiceless obstruents in syllable onset
- $F_0$ is typically 5-10 Hz higher after a voiceless obstruent as opposed to a voiced one
- in tonogenesis the $F_0$ difference is increased while the voicing contrast on the consonant is decreased and eventually lost resulting in a tonal contrast
- prevalent in East Asian languages
- example from Beijing Mandarin
- tone 1 [55] High and tone 2 [35] Rise both go back to the same level pitch category (ping) in Middle Chinese
- the category split as a function of the voicing of the onset consonant

<table>
<thead>
<tr>
<th>voiceless (yin)</th>
<th>Middle Chinese</th>
<th>Beijing Mandarin</th>
</tr>
</thead>
<tbody>
<tr>
<td>voicing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless (yin)</td>
<td>si</td>
<td>sì [55] 'poetry'</td>
</tr>
<tr>
<td></td>
<td>po</td>
<td>pò [55] 'hillside'</td>
</tr>
<tr>
<td>voiced (yang)</td>
<td>zi</td>
<td>sì [35] 'time'</td>
</tr>
<tr>
<td></td>
<td>bo</td>
<td>pò [35] 'old woman'</td>
</tr>
</tbody>
</table>

Punjabi tones (Bhatia 1975)

<table>
<thead>
<tr>
<th>Punjabi</th>
<th>Hindhi</th>
<th>Punjabi</th>
</tr>
</thead>
<tbody>
<tr>
<td>ghor-a</td>
<td>kòra</td>
<td>'horse'</td>
</tr>
<tr>
<td>dhol</td>
<td>tòl</td>
<td>'drum'</td>
</tr>
</tbody>
</table>

- voiced aspirates devoice and deaspirate but with a low tone

[8] **lexical tonal distinctions**
• **level tones**
  
  ➢ **two-way contrast:** Lingala (Bantu Congo)
    
    ```
    mo-tó 'head'  lo-lémó 'tongue'
    mo-to 'person' mo-ásí 'woman'
    mo-sísá 'vein'
    li-kolo 'leg'
    ```
  
  ➢ **three-way contrast:** Buli (Gur Ghana)
    
    ```
    syúk 'path'
    syúk 'navel'
    syúk 'fish'
    ```

• **contour tones**

  rise vs. fall Thai [Ladefoged CIP 10.4]

<table>
<thead>
<tr>
<th>high</th>
<th>mid</th>
<th>low</th>
</tr>
</thead>
<tbody>
<tr>
<td>ná: 'aunt'</td>
<td>ná: 'field'</td>
<td>ná: 'nickname'</td>
</tr>
<tr>
<td>rise ná: 'thick'</td>
<td>fall ná: 'face'</td>
<td></td>
</tr>
</tbody>
</table>

[9] **register**

• **splits pitch space into an upper and lower region**

  Cantonese [Ladefoged CIP Chapter 10 exercise G]

<table>
<thead>
<tr>
<th>high</th>
<th>mid</th>
<th>low</th>
<th>extra low</th>
<th>mid rise</th>
<th>low rise</th>
<th>high fall</th>
</tr>
</thead>
<tbody>
<tr>
<td>si</td>
<td>si</td>
<td>si</td>
<td>si</td>
<td>si</td>
<td>si</td>
<td>si</td>
</tr>
</tbody>
</table>
  | 55 'poem' | 33 'to try' | 22 'matter' | 11 'time' | 35 'to cause' | 13 'city' | 53 'silk' [
  | (for many current speakers merges with 55) |

• **just as a vowel length contrast is often supplemented with a vowel quality contrast (e.g. Arabic) so a tonal height contrast is often accompanied by a laryngeal (voice quality) contrast so that breathy voice goes with lower tone**

[9] **F0 used for intonation contours in English**

• declaratives have falling contour
• yes-no have rise
• rise also signals continuation, as in a list
• but these are just cross-linguistic tendencies; e.g. in Hungarian yes-no question ends in a fall
Figure 10.1 The relative sonority of a number of the sounds of English.