HALF-SYLLABLES IN SOUTHEAST ASIAN LANGUAGES

BECKY BUTLER
7 FEBRUARY 2013
In this talk, I will...

- investigate the notion of the half-syllable (and implicitly the sesquisyllable) in two languages – Khmer and Bunong – and its use as a diagnostic feature of Southeast Asia as a whole.
- subject the half-syllable to phonetic analysis and show that it is different in the two languages, is not a coherent linguistic unit and is therefore far from being a shared feature of Southeast Asian languages.
- use these results as a springboard to address the issue of using language as a criterion for defining geographic areas more generally.
Southeast Asia as a linguistic area

Linguistics background
  ▪ Syllables
  ▪ Speech

Phonetic experiments
  ▪ Khmer
  ▪ Bunong

Implications for half-syllables
LANGUAGES IN SOUTHEAST ASIA

- At least five different languages families
  - Mon-Khmer
  - Sino-Tibetan
  - Austronesian
  - Tai-Kadai
  - Hmong-Mien

- Over 1,000 different languages

- Remarkable internal consistency
  - vowel inventories
  - tone systems
  - minimal morphology
  - small word shapes
  - sesquisyllables
Many languages of mainland SEA are more similar to languages of neighboring China and India than they are to languages of insular SEA.

A recent conference on mainland SEA languages featured talks on Tani (India) and languages of China spoken south of the Yangtze Basin.
EFFICACY OF LANGUAGE AS A BOUNDARY MARKER WRT SOUTHEAST ASIA

- Internal consistency of language within Southeast Asia
  - Mainland/Insular

- Border between Southeast Asia and other countries/regions
  - India and China

- Division between groups within Southeast Asia
  - Upland/Lowland
SYLLABLES
SYLLABLES: UNITS OF SPEECH TIMING

- 1 Syllable
  - rice
  - com

- 2 Syllables
  - Asia
  - Cornell

- 3 Syllables
  - Washington
  - tomato

- And so on!
HALF-SYLLABLES

1 Syllable
- rice
- cōm

Sesquisyllable: One syllable preceded by a half-syllable
- σ Σ

2 Syllables
- Asia
- Cornell
- Sounds not spelling!

- English (kind of)
  - about (vs. able)
  - because (vs. beaver)

- Vowel in half-syllable is reduced: [ə]

- RELEVANCE: Supposedly very common in (mainland) Southeast Asia languages
SESQUISYLLABLES IN SOUTHEAST ASIA

- Burmese
  - [pə.lwè] ‘flute’
  - (Green 2005)

- Thai
  - [sa.nùk] ‘fun’
  - (Bennett 1994)

- Kammu (Laos)
  - [rè.háaŋ] ‘bamboo’
  - (Svantesson and Karlsson 2004)

- Moken (Burma, Thailand)
  - [phè.laː] ‘husked rice’
  - (Pittayaporn 2005)

- Kuay (Thailand)
  - [kə.thiim] ‘garlic’
  - (Preecha 1968)

- Vietnamese 😞
SPEECH AS ARTICULATION
Discrete v. Continuous

- Speech as discrete units (consonants and vowels)
  cat [kæt]

- Speech as gestures

  Tongue Movement
PROPERTIES OF SPEECH

- **Overlap**

- **Underlap**
What happens when gestures separate?

Experiment: bnick tkat
Some half-syllables:

- Real syllables: CVC
- Not real syllables: CŒC

Sesquisyllables:

- Disyllables: CVCVC
- Monosyllables: CCVC

Half-syllables don’t actually exist!
Lots of things sound like [ə]. How can we tell them apart?

<table>
<thead>
<tr>
<th>Stressed [ə]</th>
<th>Unstressed [ə]</th>
<th>Underlap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Longest</td>
<td>Shorter</td>
<td>Shortest</td>
</tr>
<tr>
<td>Neutral Position in the Mouth</td>
<td>Equal or Higher Position in the Mouth</td>
<td>Highest Position in the Mouth</td>
</tr>
</tbody>
</table>

Ex. English: “abut” [ə.bət]
- Cambodia abuts Vietnam.
Compared sesquisyllables in two Mon-Khmer languages

- **Khmer**
  - 12 mil – 13 mil speakers
  - ~65% literacy

- **Bunong (Phnong, Mnom)**
  - 50,000 speakers
  - Literacy: Low
In addition to monosyllables, disyllables and longer words, Khmer has a wealth of clusters which might be half-syllables.

<table>
<thead>
<tr>
<th>Cl1</th>
<th>C2</th>
<th>p</th>
<th>ps</th>
<th>ph</th>
<th>pr</th>
<th>pl</th>
<th>pt</th>
<th>ptf</th>
<th>pk</th>
<th>pn</th>
<th>pn</th>
<th>pη</th>
<th>p?</th>
<th>pd</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>th</td>
<td>tr</td>
<td>tl</td>
<td>tp</td>
<td></td>
<td></td>
<td>tk</td>
<td>tm</td>
<td>tn</td>
<td>tη</td>
<td>t?</td>
<td>tb</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tf</td>
<td>tfh</td>
<td>tfi</td>
<td>tfj</td>
<td>tfp</td>
<td></td>
<td></td>
<td>tfk</td>
<td>tfm</td>
<td>tfn</td>
<td>tfη</td>
<td>tf?</td>
<td>tfb</td>
<td>tfd</td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>ks</td>
<td>kh</td>
<td>kr</td>
<td>kl</td>
<td>kp</td>
<td>kt</td>
<td>ktf</td>
<td>km</td>
<td>kn</td>
<td>kn</td>
<td>kη</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td>sr</td>
<td>sl</td>
<td>sp</td>
<td>st</td>
<td></td>
<td></td>
<td>sk</td>
<td>sm</td>
<td>sn</td>
<td>sn</td>
<td>sn</td>
<td>sη</td>
<td>s?</td>
<td>sb</td>
</tr>
<tr>
<td>m</td>
<td>ms</td>
<td>mh</td>
<td>mr</td>
<td>ml</td>
<td></td>
<td></td>
<td>mn</td>
<td>mn</td>
<td>mη</td>
<td></td>
<td></td>
<td></td>
<td>m?</td>
<td></td>
</tr>
<tr>
<td>l</td>
<td>lh</td>
<td></td>
<td></td>
<td>lp</td>
<td></td>
<td></td>
<td>lk</td>
<td>lm</td>
<td></td>
<td></td>
<td></td>
<td>lη</td>
<td>l?</td>
<td>lb</td>
</tr>
</tbody>
</table>
Participants
- 18 Khmer speakers, ages 18 – 44 (μ = 27)
- Recorded at Royal University of Phnom Penh

Stimuli
- Potential sesquisyllables: 20
- Disyllables: 4
- Monosyllables: 13
Record three repetitions of supposed sesquisyllables in a frame sentence

- និឃាឃ ផ្កា ម្តង ទៀត.
  /nijij pka mdoŋ tiət/

- និឃាឃ ដ្ងៀម ម្តង ទៀត.
  /nijij sŋiəm mdoŋ tiət/
**DISTRIBUTIONAL RESULTS**

<table>
<thead>
<tr>
<th>C1 is voiced ([m] or [l])</th>
<th>C1 is voiceless ([p] or [t])</th>
</tr>
</thead>
<tbody>
<tr>
<td>Underlap in 93% of tokens</td>
<td>Underlap in 95% of tokens</td>
</tr>
<tr>
<td>Underlap is voiced 100% of the time</td>
<td>Underlap is voiceless 99% of the time</td>
</tr>
</tbody>
</table>

![Graph showing gestures and voicing differences]
Durations of [ə] and [ɐ] are **not** significantly different.

Duration of [ə] in clusters is significantly ($p<0.0001$) shorter than [ə] in unstressed syllables.

- [mteh]/[məteh] ‘pepper’ ម្ល៉េ្ ម្​េ្
- [mət.ˈpɔt] ‘stretch one’s back’ ឈឺីត់ ឈឺីត់
VOWEL POSITION

C1 – labial

+ [ə] in sesquisyllables

[ə] in disyllables and monosyllables

C1 – alveolar
- Half-syllables in Khmer are actually underlap.
- They do not have an associated gesture.
- They are not phonologically real.

"Sesquisyllables" in Khmer are monosyllables.
BUNONG
BUNONG BACKGROUND

- Not much previous work

- Not severely endangered but vulnerable

- Many recordings done at headquarters of International Cooperation Cambodia’s Research Education and Development (READ) Project headquarters in Sen Monorom, Mondulkiri

- Some recordings done in Bunong village of Bou Sra
Along with monosyllables, Bunong is also claimed to have sesquisyllabic words but no longer words.

<table>
<thead>
<tr>
<th>Monosyllables</th>
<th>Sesquisyllables</th>
</tr>
</thead>
<tbody>
<tr>
<td>[briː]</td>
<td>[rə.ʌ.laːw]</td>
</tr>
<tr>
<td>‘forest’</td>
<td>‘more than’</td>
</tr>
<tr>
<td>[pləj]</td>
<td>[lə.hat]</td>
</tr>
<tr>
<td>‘fruit’</td>
<td>‘tightly fitting’</td>
</tr>
<tr>
<td>[cuəj]</td>
<td>[kə.tojə]</td>
</tr>
<tr>
<td>‘offend’</td>
<td>‘hatchet’</td>
</tr>
<tr>
<td>[khʌt]</td>
<td></td>
</tr>
<tr>
<td>‘die’</td>
<td></td>
</tr>
<tr>
<td>[kuʔ]</td>
<td></td>
</tr>
<tr>
<td>‘sit’</td>
<td></td>
</tr>
<tr>
<td>[koːɲ]</td>
<td></td>
</tr>
<tr>
<td>‘uncle’</td>
<td></td>
</tr>
</tbody>
</table>
METHODOLOGY

- **Participants**
  - 12 Bunong speakers, ages 22 – 36 (μ = 28)
  - All men

- **Stimuli**
  - Potential sesquisyllables: 12
  - Monosyllables (CCVC): 7
  - Monosyllable (CVC): 21

- **Three repetitions in frame sentence:**
  - /lah nau kalo/ ‘above’
  - /lah nau klən/ ‘to miss’
- Bunong has far fewer consonant clusters than Khmer, but some underlap is still present
- Distribution of voicing is more variable

<table>
<thead>
<tr>
<th>C1</th>
<th>C2</th>
<th>r</th>
<th>l</th>
</tr>
</thead>
<tbody>
<tr>
<td>k</td>
<td></td>
<td>ө 98%</td>
<td></td>
</tr>
<tr>
<td>s</td>
<td></td>
<td>ө 80%</td>
<td></td>
</tr>
<tr>
<td>m</td>
<td></td>
<td>ө 89%</td>
<td></td>
</tr>
<tr>
<td>η</td>
<td></td>
<td>ө 98%</td>
<td>ө 13%</td>
</tr>
</tbody>
</table>
DURATIONS

- Durations of [ǝ] and [ə] are **not** significantly different.

- Duration of [ǝ] in clusters is significantly ($p<0.0001$) shorter than [ǝ] in unstressed syllables.

- Duration of unstressed [ǝ] is significantly shorter than stressed [ǝ] ($p<0.0001$).

- [klǝŋ]/[kǝlǝŋ] ‘to miss’ ក្លឹង
- [kǝnaːr] ‘wing’ ក្លាយ
- [kǝl] ‘big turtle’ ក្ឹុង
Underlap “vowels” are higher in the mouth than half-syllable vowels ($p < 0.0001$).
Unstressed schwa vowels (i.e. half-syllable vowels) are higher in the mouth than stressed schwa vowels.
Half-syllables in Bunong are actually syllables.
- They have an associated gesture.
- They are phonologically real.

“Sesquisyllables” in Bunong are disyllables.
RESULTS AND IMPLICATIONS
## COMPARISON OF RESULTS

<table>
<thead>
<tr>
<th>Khmer</th>
<th>Bunong</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CCVC</strong></td>
<td><strong>Cv.CVC</strong></td>
</tr>
<tr>
<td>Half-syllable is</td>
<td>Half-syllable is</td>
</tr>
<tr>
<td><strong>underlap</strong></td>
<td>a real syllable</td>
</tr>
<tr>
<td>Sesquisyllable is a</td>
<td>Sesquisyllable is a</td>
</tr>
<tr>
<td><strong>monosyllable</strong></td>
<td><strong>disyllable</strong></td>
</tr>
</tbody>
</table>
VALIDITY OF THE SESQUISYLLABLE

- Too broad
  - Subsumes two different linguistic entities – monosyllables and disyllables – which are already well-studied in the linguistic literature.

- Too narrow
  - Used as a defining characteristic of SEA, but not at all exclusive to SEA languages.

<table>
<thead>
<tr>
<th>Monosyllables with underlap</th>
<th>Disyllables with similar stress patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Georgian</td>
<td>Hixkaryana</td>
</tr>
<tr>
<td>Polish</td>
<td>Choctaw</td>
</tr>
<tr>
<td>Nuxalk (Salish)</td>
<td>Chickasaw</td>
</tr>
<tr>
<td>Tashlhiyt Berber</td>
<td>Cayuga</td>
</tr>
</tbody>
</table>
EFFICACY OF THE SESQUISYLLABLE

- Internal consistency of language within Southeast Asia
  - Mainland/Insular
    - Inconclusive; Need more data

- Border between Southeast Asia and other countries/regions
  - India and China
    - Maybe with India; probably not with southern China

- Division between groups within Southeast Asia
  - Upland/Lowland
    - No. Syllable type does not correlate with Zomia borders
      (Eg. Burmese)
EFFICACY OF LANGUAGE

- Is language useless as a boundary marker?
  - Probably not

- Are large scale generalizations useful?
  - Also probably not

- Since the sesquisyllable, which has long been considered one of the most salient defining properties of SEA languages, is a problematic concept, I suggest that more detail-oriented linguistic fieldwork is necessary to draw meaningful conclusions about the correlations between language and boundaries.
THANK YOU!