Rhythm in four Colombian Languages
Kamsá, Kogi, Embera-Chamí and Wuayuú

David Páez Acevedo
dandresp@unm.edu

Karol Ibarra Zetter
karol@unm.edu

University of New Mexico
1. Introduction
   1.1 Rhythm
   1.2 Colombian languages

2. Methodology
   Data from each language
   Measurements and Analyses

3. Results
   Deltas and PVI of sampled languages
   Comparisons with world languages

4. Discussion
Rhythm
1.1. Rhythm

- Pike (1946) and Abercrombie (1965, 1967): described classes of rhythm:
  - **Stressed-timed**: The stressed syllables are longer than unstressed ones, for instance, English.
  - **Syllable-timed**: All syllables are *isochronic*, commonly illustrated by Spanish.
1.1. Rhythm

Idealized stress timing


Idealized syllable timing

1.1. Rhythm

- These categories are problematic:
  - perceptually salient, but challenging to establish acoustic correlates.

- Grabe & Low (2002): described different methods to account for rhythm.

- Different results have been obtained.
1.1. Rhythm

- Ramus et al (1999) have shown a continuous distribution of rhythmic patterns, using:
  - $\Delta C$: standard deviation of duration of Consonantal intervals.
  - $\Delta V$: standard deviation of duration of Vowel intervals.
  - %V: proportion of vocalic intervals.
- Syllable-timed: $+V\%$, $-\Delta C$, $-\Delta V$
- Stressed-timed: $-V\%$, $+\Delta C$, $+\Delta V$
1.1. Rhythm

- Relation $\Delta V$ and $\Delta C$
  - Syllable-timed languages would present lower values of $\Delta V$ and $\Delta C$ than stress-timed languages
1.1. Rhythm

- Relation \( %V \) and \( \Delta C \)
  - higher values of \( %V \) indicate that the language is more Syllable-timed.
1.1. Rhythm

- Grabe & Low (2002) proposed measuring Pairwise Variability Indexes (PVI), which allow to account for variation in speech rate.

- **PVI** Calculates duration of a syllable relative to another
  - Raw PVI = rPVI: Consonants
  - Normalized PVI = nPVI: vowels
  - Syllable-timed: -nPVI -rPVI
  - Stressed-timed: +nPVI +rPVI
1.1. Rhythm

- PVI
  - Syllable-timed languages would present lower values of nPVI and rPVI than stress-timed languages
1.1. Rhythm

Additionally...

- It has been suggested that syllable complexity and vowel length are also related with rhythmic patterns:
  - Easterday et al (2011) found that there IS a relation $\Delta C /$ syllable complexity.
Colombian Languages
1.2. Colombian Languages

- ~64 languages in Colombia
- Limited research:
  - Grammars & Dictionaries
  - Phonological descriptions
    (mostly segmental; prosodic information limited or absent)
  - Phonetic detail scarce
  - Rhythm
    very few characterizations, mostly qualitative with no phonetic measurements (Keels 1985 for Guayaber, and Gralow 1985 for Coreguaje)
METHODOLOGY
2. Methodology: Data

- 4 Colombian languages (geographically and genetically diverse):
  - Kamsá, Kogi, Embera-Chamí and Wayuú

- Data Sources:
  - Global Recordings: Kogi, Embera-Chamí and Wayuú
  - Archive of the Indigenous Languages of Latin America (AILLA): Kamsá

- Type of data:
  - Narratives
2. Methodology: Data

- LAPSyD/WALS criteria:
  - Syllcat:
    - Complexity of the syllable: Simple, Moderate, Complex
  - Syllabic Index:
    - Maximal degree of elaboration in Onset, Nucleus and Coda (e.g., 3,1,2)
  - Canonical Form
    - E.g., (C)V(C)
2. Methodology: Data

- WAYUÚ
- KOGI
- EMBERA-CHAMÍ
- KAMSA
2. Methodology: Data

- Classification: Arawakan, Maipuran, Northern
- Location: La Guajira peninsula on Venezuela/Colombia border
- Syllcat: Moderate
- Syllabic index: 2, ONC: 011
- Canonical Form: (C)V(C)

Source: LAPSyD
2. Methodology: Data

15 Consonants
6 Vowels

Source: LAPSyD
2. Methodology: Data

- Classification: Chibchan
- Location: Sierra Nevada de Santa Marta, Colombia
- Syllcat: Moderate
- Syllabic index: 2
- Canonical Form: (C)V(V)(C)

Source: LAPSyD
2. Methodology: Data

15 Consonants
7 Vowels

Source: LAPSyD

<table>
<thead>
<tr>
<th>Vowels</th>
<th>front unrounded</th>
<th>central unrounded</th>
<th>back unrounded</th>
<th>back rounded</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>i</td>
<td>i</td>
<td>u</td>
<td></td>
</tr>
<tr>
<td>higher mid</td>
<td>e</td>
<td></td>
<td>o</td>
<td></td>
</tr>
<tr>
<td>raised low</td>
<td>a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Consonants</th>
<th>bilabial</th>
<th>alv pal-</th>
<th>alv palata</th>
<th>velar</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>stop</td>
<td>p</td>
<td>t</td>
<td>k</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>fric pulm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nasal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>appr lat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: KOGI
2. Methodology: Data

- Classification: Chocó
- Location: Departments of Risaralda, Caldas, Antioquia, Valle, Colombia
- Syllcat: Moderate
- Syllabic index: 3  
  ONC: 111
- Canonical Form: C(C)V(C)

Source: LAPSyD
### 2. Methodology: Data

**EMBERA-CHAMÍ**

#### 18 Consonants

<table>
<thead>
<tr>
<th>Consonants</th>
<th>Bilabial</th>
<th>lab-den</th>
<th>alv-pal</th>
<th>alv-palatal</th>
<th>velar</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>stop</td>
<td>-vless</td>
<td>-vless</td>
<td>-vless</td>
<td></td>
<td>p</td>
<td>t</td>
</tr>
<tr>
<td>aff</td>
<td>sib</td>
<td>-vless</td>
<td>-vless</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fric</td>
<td>sib</td>
<td>-vless</td>
<td>-vless</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pulm</td>
<td>sib</td>
<td>-vless</td>
<td>-vless</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nasal</td>
<td>-vless</td>
<td>-vless</td>
<td>-vless</td>
<td></td>
<td>m</td>
<td>n</td>
</tr>
<tr>
<td>trill</td>
<td>-vless</td>
<td>-vless</td>
<td>-vless</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tap/flap</td>
<td>-vless</td>
<td>-vless</td>
<td>-vless</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>appr</td>
<td>-vless</td>
<td>-vless</td>
<td>-vless</td>
<td></td>
<td>w</td>
<td>v</td>
</tr>
<tr>
<td>impl stop</td>
<td>-vless</td>
<td>-vless</td>
<td>-vless</td>
<td></td>
<td>b</td>
<td>d</td>
</tr>
</tbody>
</table>

#### 12 Vowels

<table>
<thead>
<tr>
<th>Vowels</th>
<th>front unrounded</th>
<th>central unrounded</th>
<th>back unrounded</th>
<th>rounded</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>i</td>
<td>u</td>
<td>u</td>
<td></td>
</tr>
<tr>
<td>nasalized</td>
<td>i</td>
<td>ü</td>
<td>ü</td>
<td></td>
</tr>
<tr>
<td>higher mid</td>
<td>e</td>
<td>ø</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nasalized</td>
<td>e</td>
<td>õ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>low</td>
<td>a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nasalized</td>
<td>a</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: LAPSyD
2. Methodology: Data

- Classification: Isolate
- Location: Western Colombia, Sibundoy Valley
- Syllcat: Complex
- Syllabic index: 6* ONC: 231
- Canonical Form: C(C)(C)V(V)(V)(C)

Source: LIC/WALS
2. Methodology: Data

### Consonants

- 22 Consonants
- 6 Vowels

Source: Lenguas de Colombia.gov

#### Consonantes

<table>
<thead>
<tr>
<th></th>
<th>Labiales</th>
<th>Alveolares</th>
<th>Retroflejas</th>
<th>Palatales</th>
<th>Velares</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oclusivas</td>
<td>pb</td>
<td>t d</td>
<td></td>
<td>k g</td>
<td></td>
</tr>
<tr>
<td>Africadas</td>
<td>ts tś tš</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricativas</td>
<td>p s š</td>
<td>š x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasales</td>
<td>m n</td>
<td></td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laterales</td>
<td>l</td>
<td></td>
<td></td>
<td>y</td>
<td></td>
</tr>
<tr>
<td>Vibrantes</td>
<td>ř</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semivocales</td>
<td>w</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Vocales

<table>
<thead>
<tr>
<th></th>
<th>Anteriores</th>
<th>Centrales</th>
<th>Posteriores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altas</td>
<td>i</td>
<td>i</td>
<td>u</td>
</tr>
<tr>
<td>Bajas</td>
<td>e</td>
<td>a</td>
<td>o</td>
</tr>
</tbody>
</table>
2. Methodology: Measurements

- 2 minutes of speech
- Vocalic and consonantantal intervals in continuous speech
  - a sequence of segments (V or C) is an interval regardless of word limits: [VV#V] ~ [C.CC]
- Duration of the interval in Praat
- Initial and Final segments in IU were ignored
2. Methodology: Measurements
RESULTS
## 3. Results

<table>
<thead>
<tr>
<th>Language</th>
<th>Structure</th>
<th>rPVI</th>
<th>nPVI</th>
<th>% V</th>
<th>ΔV</th>
<th>ΔC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wayuú</td>
<td>CVC = 2</td>
<td>65.2</td>
<td>55.1</td>
<td>50.6</td>
<td>69.9</td>
<td>65.1</td>
</tr>
<tr>
<td>Kogi</td>
<td>CVVC = 3</td>
<td>65</td>
<td>52.2</td>
<td>48.4</td>
<td>68.3</td>
<td>77.6</td>
</tr>
<tr>
<td>Embera-chamí</td>
<td>CCVC = 3</td>
<td>39.7</td>
<td>46.6</td>
<td>50.5</td>
<td>52.3</td>
<td>37.6</td>
</tr>
<tr>
<td>Kamsá</td>
<td>CCCVVVC = 6</td>
<td>73.7</td>
<td>47.1</td>
<td>49.7</td>
<td>57.3</td>
<td>69.3</td>
</tr>
</tbody>
</table>
3. Results: $\Delta C$, %V

- Embera: Syllable timed
- Kogi: Stress timed
Compared to Easterday et al. (2011).

- Embera: Syllable timed
- K,W,K: closer to each other
Compared to Romano & Mairano (2010).

K, W, K: closer to Arabic and Tedesco, but rather separated from English, a typically known stress-timed language.

Embera: very far from Spanish; highly syllable timed
3. Results: $\Delta C, \Delta V$

- Embera: Highly Stress timed
- Kogi and Wayuú, Highly Syllable timed
- Kamsá: in the middle
3. Results: $\Delta C, \Delta V$

- Compared to Romano & Mairano (2010)
- K,W,K: Extremely distant
- Emberá: Closer, but still far from Spanish.
3. Results: PVI

- Embera: Highly Syllable timed
- Kogi and Wayúu: Stressed timed
- Kamsá:
  - lower values of rPVI = more stressed timed
  - Higher values of nPVI = more syllable timed
3. Results: PVI

- Compared to Grabe & Low (2002)
- Embera: close to Syllable timed languages.
- Kogi and Wayúú, close to Stress timed languages
- Kamsá: close to Polish (A mixed language!)
DISCUSSION
4. Discussion

- The analyses of Deltas and PVI values revealed a pattern according to which:
  - Embera: syllable-timed
  - Kogi and Wayuu: stress-timed
  - Kamsá: mixed language (apparently)

- Kamsá reportedly with the most complex patterns of in syllabic structure.

- Possibly a reason for Kamsá’s distinct behavior.
4. Discussion

- All four Colombian languages are distributed farther in the continuum of $\Delta C$ and $\% V$ from the more widely known languages (Romano & Mairano 2010). In that sense, these languages can serve as points of reference on the continuum.

- Nevertheless, comparison with less known languages as those reported by Easterday et al (2011), and the more common ones, reported by Grabe & Low (2002), the Colombian languages are in the middle of the continuum.
4. Discussion: Future research

- Correlation between syllable complexity, vowel length, vowel reduction and rhythm.

- There is a possibility that this correlation is central in the consideration of Kamsá as a mixed language.
4. Discussion: Future research

- It is important to run perceptual studies that corroborate that our results correspond to actual perceptual experience.
References (a)

References (b)

Rhythm in four Colombian Languages
Kamsá, Kogi, Embera-Chamí and Wuayúú

David Páez Acevedo
dandresp@unm.edu

Karol Ibarra Zetter
karol@unm.edu

University of New Mexico

Gracias
REFERENCES:

- Emberá-Chamí. LAPSyD

- Kamsa AILLA

- Kogi LAPSyD

- Wayuú LAPSyD
References

- Kogi: Audio Resources - Global Recordings Network
  http://globalrecordings.net/en/language/240

- Emberá-Chocó: Crandell, Rachel. Global Recordings Network:
  http://globalrecordings.net/en/language/3499

- Kamsa : Lenguas Indígenas de Colombia (LIC).

- Wayuu Global Recordings Network:
  http://globalrecordings.net/program/Co4230.