The analysis of complex tonal systems: motivations, methods, and analysis

Dr. Christian DiCanio
cdicanio@buffalo.edu

University at Buffalo

5/29/18
Meta-outline for the lectures

1. The analysis of complex tonal systems: motivations, methods, and analysis
2. Speech perception in the field
3. Creating and working with endangered language corpora
4. Higher-level prosody and tone
Goals

1. Explore complex prosodic systems
2. Examine methods for working with and analysing complex tonal systems
3. Present case studies illustrating these approaches

Please feel free to ask questions.
## Types of prosodic systems

<table>
<thead>
<tr>
<th>Tonal</th>
<th>Non-tonal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical stress</td>
<td>English, Spanish, Polish, Mixe</td>
</tr>
<tr>
<td>Triqui, Ixcatec, Thai,</td>
<td></td>
</tr>
<tr>
<td>Serbo-Croatian</td>
<td></td>
</tr>
<tr>
<td>Cantonese, Vietnamese,</td>
<td>French, Hindi, Cayuga</td>
</tr>
<tr>
<td>Mambila</td>
<td></td>
</tr>
</tbody>
</table>

Many tone languages with stress have fixed stress, but fixed stress is also not uncommon among non-tonal languages.

Languages with neither lexical stress nor tone might have complex rules for phrase-level accentuation.

Complex tonal systems

A complex prosodic system is a language possessing any one of the following characteristics:

- The language has > 3 tonal contrasts (c.f. Maddieson (2010)).
- Both tone and stress interact.
- Tone and phonation type interact (but one is not strictly predictable based on the other).
- Tone and length interact.
Where are the complex prosodic systems?

- Oto-Manguean family - Southern Mexico
- Atlantic-Congo family - Sub-Saharan Africa
- Sino-Tibetan family - East/Southeast Asia
- Austroasiatic family - Southeast Asia
- Tai-Kadai family - Southeast Asia
- Khoisan family - Southern Africa
- Central Sudanic languages
- Trans New Guinea languages
Distribution of tone languages

(Maddieson, 2011)
**Example - Tone in Itunyoso Triqui**
DiCanio (2014)

Table: Tones on Monosyllables by rime type

<table>
<thead>
<tr>
<th>Tone</th>
<th>CV</th>
<th>Gloss</th>
<th>CV?</th>
<th>Gloss</th>
<th>CVh</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>/4/</td>
<td>β:e 4</td>
<td>hair</td>
<td>n:i? 4</td>
<td>our mother</td>
<td>β:e h 4</td>
<td>beat.3sg (intr.)</td>
</tr>
<tr>
<td>/3/</td>
<td>n:e 3</td>
<td>plow</td>
<td>n:e? 3</td>
<td>straw rope</td>
<td>n:e h 3</td>
<td>dream</td>
</tr>
<tr>
<td>/2/</td>
<td>n:e 2</td>
<td>to lie (tr.)</td>
<td>n:i? 2</td>
<td>smelly</td>
<td>β:e h 2</td>
<td>cave</td>
</tr>
<tr>
<td>/1/</td>
<td>n:e 1</td>
<td>naked</td>
<td>?ni? 1</td>
<td>be.salty</td>
<td>cnãh 1</td>
<td>brother (voc.)</td>
</tr>
<tr>
<td>/45/</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>β:e h 35</td>
<td>straw mat</td>
</tr>
<tr>
<td>/13/</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>keh 13</td>
<td>barely</td>
</tr>
<tr>
<td>/43/</td>
<td>li 43</td>
<td>small</td>
<td>*</td>
<td>*</td>
<td>n:ãh 43</td>
<td>mother (voc.)</td>
</tr>
<tr>
<td>/32/</td>
<td>n:e 32</td>
<td>water</td>
<td>*</td>
<td>*</td>
<td>n:ãh 32</td>
<td>cigarette</td>
</tr>
<tr>
<td>/31/</td>
<td>n:e 31</td>
<td>meat</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>
Why study an endangered language’s sound system?

- Exploratory: I want to understand how the language works.
- Typological: This language has features not found/examined in other languages.
- Specific research motivation, e.g. Tone should align to syllables, not moras (c.f. Xu (1999)).

“We often forget how science and engineering function. Ideas come from previous exploration more often than from lightning strokes... Finding the question is often more important than finding the answer.” (Tukey, 1980).
The role of the phonetician/linguist

The short-term phonetician - you know the contrasts and want to do a quick study on some aspect of the language

The long-term phonetician - you have worked on the language and want to delve into a deeper question

The corpus data phonetician - you are working on existing recordings of the language to answer phonetic questions

***All of these roles assume that you have established the tonal/prosodic contrasts on words in the language. If not, you need to.
Establishing tonal contrasts

Let’s assume that you are just starting to do fieldwork and you think the language is tonal.

It is crucial to identify the contrasts before phonetic analysis - you can not just “look at Praat and find out what the tones are.”

Various instruments which record speech—for example, the phonograph, the dictaphone, and magnetic-wire or magnetic-tape recorders—can be of considerable aid to the investigator, since they permit the repetition of phrases. The use of such machines, however, has two grave dangers: (1) The investigator tends to deprive himself of hearing the natural range of key and free variation which comes in repetition by the informant and may, therefore, record as different some utterances of tonemes which are functionally the same in spite of temporary slight, free, pitch divergences. Sufficient recordings of repetitions by the informant himself would overcome this danger. (2) The investigator is tempted to be too "accurate," that is, to transcribe (just because he can find them with instruments) details which do not reflect the system, but are changes within tonemes. Here the danger can be avoided if the investigator uses such data to describe the tonal variants but for publication of grammatical and phonetic studies uses a written transcription which records only the significant tone units (tonemes).

(Pike, 1948:43)
Pike provides a framework for the analysis of tone languages (see also Yu (2014)):

1. Sort words into similar sized units, e.g. all disyllabic words.
2. Pick two words and either (a) ask your consultant to produce them or (b) play the recordings
3. Is the F0 pattern is the same or different?
4. If different, place the words into separate bins. If same, place the words into the same bin.
5. Take the next word and compare it to the previous two.
6. Change the bin of a word as needed.
7. The name of each bin is arbitrary at first. Decide on this later.

–exercise using Toney–
What is an Otomanguean language?

180 languages spoken in Southern Mexico
Tonal complexity in Otomanguean languages

Tonal inventory size

- 2 - 3
- 4 - 5
- 6 - 7
- 8 - 9
- 10 - 11

Made with Google My Maps
All are tonal and many have very large tonal inventories. At least three tones are reconstructed at the earliest levels (Kaufman, 1990; Rensch, 1976).

Laryngeal/glottal features which are often orthogonal to tone (Silverman, 1997a).

Complex onsets are possible, but most languages lack codas. Most languages have polysyllabic words (DiCanio and Bennett, 2018).

Complex morphology on verbs and with personal clitics which frequently involves tone (Campbell et al., 1986; DiCanio, 2016; Palancar, 2009; Suárez, 1983) and classic processes of tone sandhi (Goldsmith, 1990; Pike, 1948).
## Complex tonal systems

How many tones occur in Oto-Manguean languages?

<table>
<thead>
<tr>
<th>Language</th>
<th>Tones</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern Pame</td>
<td>2</td>
<td>(Berthiaume, 2004)</td>
</tr>
<tr>
<td>Mazahua</td>
<td>4</td>
<td>(Knapp Ring, 2008)</td>
</tr>
<tr>
<td>Tlacoatzintepec Chinantec</td>
<td>7</td>
<td>(Thalin, 1980)</td>
</tr>
<tr>
<td>Itunyoso Triqui</td>
<td>9</td>
<td>(DiCanio, 2008)</td>
</tr>
<tr>
<td>Yoloxóchitl Mixtec</td>
<td>10</td>
<td>(DiCanio et al., 2012)</td>
</tr>
<tr>
<td>San Juan Quiahije Chatino</td>
<td>11</td>
<td>(Cruz, 2011)</td>
</tr>
<tr>
<td>Chiquihuitlan Mazatec</td>
<td>17</td>
<td>(Jamieson, 1977)</td>
</tr>
<tr>
<td>Quiotepec Chinantec</td>
<td>19+</td>
<td>(Castillo Martínez, 2011)</td>
</tr>
</tbody>
</table>

But how do you count? Is the TBU the stem? the syllable? the mora?
Quetzalapa Chinantec

Five tone levels with contours (rising tones excluded). Words courtesy of Isabel Alhondra.

<table>
<thead>
<tr>
<th>Tone</th>
<th>Word</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>55</td>
<td>tso'</td>
<td>‘his/her fault’</td>
</tr>
<tr>
<td>44</td>
<td>tso'</td>
<td>‘illness’</td>
</tr>
<tr>
<td>33</td>
<td>tso'</td>
<td>‘he/she goes’</td>
</tr>
<tr>
<td>22</td>
<td>tso'</td>
<td>‘straight’</td>
</tr>
<tr>
<td>21</td>
<td>tso'</td>
<td>‘sin’</td>
</tr>
<tr>
<td>32</td>
<td>tso'</td>
<td>‘male’</td>
</tr>
<tr>
<td>42</td>
<td>tso'</td>
<td>‘people’</td>
</tr>
</tbody>
</table>

What is the TBU here though? Are there only 5 levels (1/mora)?
Fieldwork on Oto-Manguean languages

1. Early work during the colonial period on Mixtec grammar (de Alvarado, 1593; de Los Reyes, 1593) and Zapotec grammar (de Córdoba, 1578).

2. Comparative work on different Oto-Manguean languages in the early years of the republic (Belmar, 1897; León, 1902; Mechling, 1912; IHS, 1893).

3. Substantial SIL research on many different Oto-Manguean languages between the early 1930’s and the late 1970’s (Pike and Pike, 1947; Longacre, 1952; Mak, 1953).

Diversity within each “language”

Despite more research on these languages, there are few instrumental phonetic studies. Moreover, there is extensive diversity even within a single “language”, e.g. Copala, Chicahuaxtla, and Itunyoso Triqui examples below (DiCanio, 2008; Hollenbach, 1984; Hernández Mendoza, 2017).

<table>
<thead>
<tr>
<th>Itunyoso</th>
<th>Chicahuaxtla</th>
<th>Copala</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>ra³ʔa³</td>
<td>ra³ʔa³</td>
<td>ra³ʔa³</td>
<td>‘hand’</td>
</tr>
<tr>
<td>cnãh³5</td>
<td>ði⁴nì⁴</td>
<td>ti³nũ³5</td>
<td>‘my brother’</td>
</tr>
<tr>
<td>τʃa³tã³</td>
<td>jã³tã³</td>
<td>τʃa³tã³</td>
<td>‘pineapple’</td>
</tr>
<tr>
<td>τʃa³kah⁵</td>
<td>jã³ka⁵</td>
<td>jka³⁵</td>
<td>‘pig’</td>
</tr>
<tr>
<td>ßßeh³</td>
<td>wweh(e)³</td>
<td>ju³ßeh³</td>
<td>‘boundary stone’</td>
</tr>
<tr>
<td>tṯsoh³</td>
<td>ṯsoh(o)³</td>
<td>nì³ṯsoh³</td>
<td>‘female’s belt’</td>
</tr>
</tbody>
</table>

The distance between each town is approximately 5.2 miles.
Itunyoso Triqui (Oto-Manguean: Mixtecan)

- One of 3 dialects (Itunyoso, Chicahuaxtla, Copala); spoken in San Martín Itunyoso and La Concepción Itunyoso
- 2500 speakers; mostly bilingual but Triqui-dominant
Evolution of fieldwork

1. Invitation to write a dictionary.

2. Initial stages of exploration/fieldwork; learning the sounds.

3. Phonetic studies investigating speech production; tone production.

4. Speech perception research.

5. Language documentation/grammatical research.
The initial stages of fieldwork involve exploring the sound system and basic grammar of the language. One type of topic for production research is an investigation of how two phonological contrasts interact.

e.g. tone and glottalization; stress and word size; tone and stress, etc.

While the initial stages of this work are exploratory, in relation to suprasegmental contrasts there are a number of active research areas:

- **Glottalization and tone production** (Andruski, 2006; DiCanio, 2009; Esposito, 2010; Garellek and Keating, 2011)

- **Tonal coarticulation** (Chang and Hsieh, 2012; Gandour et al., 1994; Shen, 1990; Zhang and Liu, 2011)

- **Tonal alignment** (Arvaniti et al., 2006; Mücke et al., 2009; Xu, 1998, 1999; Xu and Liu, 2006; Zsiga and Nitisaroj, 2007)
Triqui phonology

\[ na^2-ki^3-\tilde{y}ah^3 \quad sa^1=sih^3 \quad y\tilde{a}^3\tilde{a}^3 \]

CAUS-PERF-make good=3SM light
‘He is fixing the light.’

(DiCanio, 2016:227)

- Most words are disyllabic (63% of the lexicon), many are monosyllabic (27%), and some (10%) are trisyllabic. CV, CV?/h syllable structure with possible stem-initial clusters (s/r+stop).

- Stem-final syllables are stressed. Prenasalized stops, glottalized sonorants, nasal vowels, glottal consonants, and the full tonal inventory only surface on stem-final syllables (DiCanio, 2008).

- Like many Oto-Manguean languages, the morphemic inventory is small (around 1600 roots). There are many lexicalized compounds.
Triqui experiment - (DiCanio, 2012)

<table>
<thead>
<tr>
<th>Tone</th>
<th>Open σ</th>
<th>Coda /h/</th>
<th>Coda /ʔ/</th>
<th>/V?V(h)/</th>
</tr>
</thead>
<tbody>
<tr>
<td>/4/</td>
<td>bêe⁴</td>
<td>yāh⁴</td>
<td>tfi⁴</td>
<td>ra⁴ʔa⁴</td>
</tr>
<tr>
<td>/3/</td>
<td>nne³</td>
<td>yāh³</td>
<td>tsi³</td>
<td>na³ʔa³</td>
</tr>
<tr>
<td>/2/</td>
<td>nne²</td>
<td>nāh²</td>
<td>ttfi²</td>
<td>ta²ʔa²</td>
</tr>
<tr>
<td>/1/</td>
<td>nne¹</td>
<td>kāh¹</td>
<td>tfi¹</td>
<td>na¹ʔa¹</td>
</tr>
<tr>
<td>/45/</td>
<td>bêi¹³</td>
<td>nāfi⁴⁵</td>
<td>to wash</td>
<td>nā³ʔa³⁵</td>
</tr>
<tr>
<td>/13/</td>
<td>bêi¹³</td>
<td>nāh¹³</td>
<td>this (one)</td>
<td>kā¹ʔa³</td>
</tr>
<tr>
<td>/43/</td>
<td>tʃe⁴³</td>
<td>nnāh⁴³</td>
<td>‘mother! (voc.)’</td>
<td>ko⁴ʔo⁴³</td>
</tr>
<tr>
<td>/32/</td>
<td>nne³²</td>
<td>nnāh³²</td>
<td>‘cigarette’</td>
<td>sā³ʔa²⁴</td>
</tr>
<tr>
<td>/31/</td>
<td>nne³¹</td>
<td>‘meat’</td>
<td></td>
<td>kā³ʔa¹</td>
</tr>
</tbody>
</table>

Itunyoso Triqui has 9 tones and a contrast between open syllables and those with final glottal consonants /ʔ, h/. There is also an intervocalic /ʔ/. How do speakers distinguish these tones in production when we know that non-modal phonation type can influence F₀?

How do speakers distinguish these tones in production when we know that non-modal phonation type can influence F₀?
Laryngeal consonants and non-modal phonation types induce pitch perturbations on adjacent vowels.

- Creaky and breathy phonation cause $F_0$ lowering (Garellek, 2012; Garellek and Keating, 2011; Gordon and Ladefoged, 2001; Keating et al., 2015; Wayland and Jongman, 2003).


- Restricts the phonological distribution and phonetic phasing of laryngeals with respect to tone in languages of the world (Silverman, 1997a,b).
Tone-Laryngeal Phasing (Silverman, 1997b)

- Laryngeals will perturb pitch, causing lowering or raising. This causes the listener to misperceive the tonal target.

- **Abrupt phasing** will restrict the degree of coarticulation between the laryngeal pitch perturbation and the tonal target.

- **Gradual phasing** will permit laryngeal pitch perturbation to overlap with the adjacent vowel’s tone.

- Abrupt phasing permits the listener to *recover* the intended tonal target on the vowel. Gradual phasing does not.

- If tone-laryngeal coarticulation is minimized to prevent possible listener misperception, one expects abrupt phasing.
Overlap interferes with tonal cues.

Non-overlap permits tone to be more clearly cued.
Prediction: In languages with large tonal inventories, we expect abrupt phasing.

Figure 1: Abrupt Phasing, [tt\textipa{[i]}\textipa{[2]}] ‘ten’
Figure 2: Gradual Phasing, [ββeh₄⁵] ‘straw mat’
Where are we in the research pipeline?

1. We know the contrasts and we know a potential interaction.
2. We have identified a particular research question relevant to the contrasts.

But how can we investigate this question? What methodological issues arise?
What we can and can not do

We can select a set of words which contrast each of the tones which occur with glottalization.

We can use a natural carrier sentence, /ni^[4][ja][43] Target nã^[3] / ‘I see Target here.’

But tone is also unevenly distributed. There are very few lower tone nouns/verbs and very few higher tone adjectives.

Given this distribution (and the lexicon), it will be difficult to control for onset voicing (Hombert et al., 1979; Kingston et al., 2008), nor for vowel quality (Whalen and Levitt, 1995), both of which influence F₀.
Tonal distribution by part of speech

Proportion of tones within each PoS

Adj

Verb

Noun

Adv
Method

Prediction: Laryngeal consonants will be abruptly phased with adjacent vowels to ensure listener perception of tonal contrast.

- Acoustic recordings of words in natural carrier sentences, but differed for adjectives/verbs than from nouns.
- Stimuli: 82 words per subject (5 repetitions), comprising all tonal and laryngeal conditions.
- Laryngeal conditions: VV, Vḫ, Vʔ, VʔVḫ.
- 6 speakers (4 male, 2 female), native speakers from San Martín Itunyoso, Oaxaca, Mexico.
- Measured pitch and spectral tilt (H1-H2, H1-A3) over final rime or /VʔVḫ/ duration with Praat scripts.
- Comparison of pitch and spectral tilt differences across tonal and laryngeal conditions.
- All data examined using linear mixed effects modeling.
Spectral tilt

The vocal folds come together more quickly in the production of creaky or tense phonation. As a result of this, vocal fold closure produces a higher amplitude transient. This excites a greater number of harmonics and resonances across the spectrum, resulting in a flatter overall spectral tilt than one finds for modal voice (Kirk et al., 1993).

By contrast, the closing phase of the vocal folds is slower for breathy phonation. The resulting weaker transients fail to excite higher resonances in the vocal tract and the spectrum is steeper.
Involves the estimation of amplitude difference between different harmonics/formants in the spectrum.
Phasing: /Vʔ/

- *Abrupt* phasing of glottalization on /Vʔ/ rimes with some creaky phonation on the preceding vowel.

**Figure 3:** [nneʔ³] ‘mecate’
Results

$F_0$ data - /Vː/ vs. /Vʔ/ rimes

Abrupt timing; small influence of glottal stop on $F_0$ of tone (for tone /3/).
Spectral tilt data - /V:/ vs. /Vʔ/ rimes

Lower spectral tilt occurs across the vowel preceding the glottal stop, indicating early glottal constriction. Though, this was only significant for the H1-A3 measure.
**F₀ data - \[/V:\]/ vs. \[/Vh]/ rimes**

Gradual timing; breathiness causes gradual decrease for all but the lowest tones. Some tones raise before the breathiness (compensation).
Spectral tilt data - /V:/ vs. /Vɨ/ rimes

Substantial variation in H1-H2 values, but lower H1-H2 during breathiness.
Spectral tilt data - /V:/ vs. /Vh/ rimes

Breathiness causes raising of H1-A3, but as $F_0$ falls, so does H1-A3.
Gradual phasing of /ʔ/ and /h/, where the former is realized as creak and the latter as breathy phonation.

Figure 5: [sõ³ʔõ²] ‘money’
**F₀ data: /Vh/ vs. /VʔVh/**

Similar final F₀ values, but creaky phonation causes F₀ lowering in the first half of the complex sequence.

“r” = “phonetically rearticulated vowel”
Spectral Tilt: /Vh/ vs. /VʔVh/

Lower H1-H2 values during creaky phonation.

“r” = “phonetically rearticulated vowel”
Discussion: Phasing

- Abrupt phasing of coda /ʔ/ with respect to vowel causes less significant pitch perturbation than observed for medial /ʔ/.
- Greater coarticulation (gradual phasing) of medial /ʔ/ causes more significant pitch perturbation.
- Greater coarticulation (gradual phasing) of coda /h/ causes significant pitch perturbations.

Conclusion: Mostly confirmatory - phasing of glottalization (measured by spectral tilt) determines degree of \( F_0 \) perturbation on tones.
Recap

- Overview of complex tonal systems from a typological perspective
- Methods for the initial analysis of a tonal system
- Oto-Manguean tonal systems
- Case study on Triqui; moving from an initial description to research on speech production.


