Triqui tonal coarticulation and contrast preservation in tonal phonology

Christian DiCanio
dicanio@haskins.yale.edu

Haskins Laboratories & Yale University
http://linguistics.berkeley.edu/~dicanio

4/6/14
SSMCA
Tone production and variability

Tone production varies as a function of many factors.

Tones differ phonetically with context - but how?
What constrains the production of a sequence of lexical tones?

1. **Tonal context**: the production of a tone is sensitive to the $F_0$ height and slope of adjacent tonal targets.

2. **Speech rate**: rate impinges on the temporal demands for producing certain $F_0$ shapes.

3. **Stress**: stressed and unstressed syllables condition differences in tone production.
What constrains the production of a sequence of lexical tones?

1. **Tonal context**: the production of a tone is sensitive to the $F_0$ height and slope of adjacent tonal targets.

2. **Speech rate**: rate impinges on the temporal demands for producing certain $F_0$ shapes.

3. **Stress**: stressed and unstressed syllables condition differences in tone production.
What constrains the production of a sequence of lexical tones?

1. **Tonal context**: the production of a tone is sensitive to the F₀ height and slope of adjacent tonal targets.

2. **Speech rate**: rate impinges on the temporal demands for producing certain F₀ shapes.

3. **Stress**: stressed and unstressed syllables condition differences in tone production.
Introduction

Goals

The goal here is to investigate some of these factors.

In other words, if we can make it \textit{really} hard for Triqui speakers to produce their tones accurately, we might start to see which dimensions exert stronger influences on tone production.
Roadmap

1. Background on Itunyoso Triqui and tonal coarticulation
2. Experiment: Effect of speech rate and context on tonal coarticulation
   - Hypotheses & Methods
   - Coarticulatory results
   - Speech rate results
   - Duration results
3. Discussion
4. Conclusions & future directions
Oto-Manguean languages

There are roughly 177 Oto-Manguean languages, all of which are tonal. Many have large tonal inventories.
Tones in Oto-Manguean: a typological abundance

- Approximately 41.8% of the world’s languages (220/527) are tonal (Maddieson, 2011).

- Of these, 60% (132/220) have only 1-2 lexical tone contrasts and 40% have three or more tonal contrasts (88/220).

- Among the tone languages with large inventories, languages with between 3-6 tonal contrasts are relatively common, e.g. Thai (5), Mandarin (4), Vietnamese (6), Cantonese (6), Yoruba (3).

- Languages with greater than 6 tones are rarer, but many are Oto-Manguean, e.g. Itunyoso Triqui (9) (DiCanio, 2008), Yoloxóchitl Mixtec (10/mora) (DiCanio et al., 2012a), Quiahije Chatino (15) (Cruz et al., today), Tlacoatzintepec Chinantec (7) (Thalin, 1980), Chiquihuitlan Mazatec (17) (Jamieson, 1977).
Spoken in the town of San Martín Itunyoso, in Oaxaca, Mexico. It is one of three Triqui variants.

Original fieldwork since 2004, with a focus on the phonetics and phonology, both from a descriptive perspective and from an experimental perspective.
Tone in Itunyoso Triqui

- Nine lexical tones contrast in word-final syllables, but only level tones occur in non-final syllables. Syllable structure is open with the exception of two possible laryngeal codas (DiCanio, 2008).
- There are morphological tone changes, but no tone sandhi across words (DiCanio, date).

<table>
<thead>
<tr>
<th>Tone</th>
<th>IPA</th>
<th>Gloss</th>
<th>Tone</th>
<th>IPA</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>β:e⁴</td>
<td>‘hair’</td>
<td>43</td>
<td>li⁴³</td>
<td>‘small’</td>
</tr>
<tr>
<td>3</td>
<td>n:e³</td>
<td>‘plough’</td>
<td>32</td>
<td>n:e³²</td>
<td>‘water’</td>
</tr>
<tr>
<td>2</td>
<td>n:e²</td>
<td>‘to lie (tr.)’</td>
<td>31</td>
<td>n:e³¹</td>
<td>‘meat’</td>
</tr>
<tr>
<td>1</td>
<td>n:e¹</td>
<td>‘naked’</td>
<td>45</td>
<td>joh⁴⁵</td>
<td>‘my forehead’</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>13</td>
<td>jo¹³</td>
<td>‘light, quick’</td>
</tr>
</tbody>
</table>
Tones in open syllables, from (DiCanio, 2012)
Stress in Triqui

- In addition to tone, Itunyoso Triqui has word-final stress.

- Stress is realized phonologically by distributional differences. Nasal vowels, laryngeals, prenasalized stops, and contour tones all surface only in word-final syllables. A reduced inventory of contrasts occurs in non-final syllables.

- Final syllables are also phonetically longer than non-final syllables (DiCanio, 2010).

![Graph showing the duration of syllables with C1, V1, C2, V2 labels and a duration scale from 0 to 500 ms.](image)
Example: /tʃu³ ku³/ ‘animal’

Note that the tone is held constant here across syllables.
Tonal context and compatibility

- Phonetic contexts which cause more abrupt $F_0$ transitions between adjacent syllables are more likely to perturb $F_0$ than contexts causing gradual transitions (Xu, 1994).

- **Conflicting:** the tone offset does not match the following onset, e.g. Rise + Low.

- **Compatible:** the tone offset matches the following tone onset, e.g. Fall + Rise.

- (Related) Contour tones may also be more susceptible to contextual variation than level tones as they have multiple phonetic targets that must be achieved within a shorter time frame.
Greater coarticulatory effects occur in conflicting contexts than in compatible contexts in Mandarin Chinese (Xu, 1994).
The effect of speech rate

- As rate increases, fewer F₀ targets are successfully reached and contours flatten (contour simplification).

- Increased rate causes a reduction in a speaker’s F₀ range in French (Fougeron and Jun, 1998).

- Clear speech has a larger F₀ range and a slower speech rate than conversational speech in English and Croatian (Bradlow et al., 2003; Smiljanić and Bradlow, 2005).

- Stress also causes contour simplification of tones in Thai (Potisuk et al., 1996).
Directionality and convergence

- Languages differ in the direction in which tonal coarticulation spreads. While all languages have some anticipatory and carryover effects, the latter is typically stronger (in Vietnamese, Thai, Mandarin) (Brunelle, 2009; Gandour et al., 1994; Xu, 1997).

- In Malaysian Hokkien, both anticipatory and carryover tonal coarticulation occur, but the former is stronger, contra findings in the previous literature (Chang and Hsieh, 2012).

- Generally, tonal coarticulation is assimilatory in nature, but anticipatory coarticulation is sometimes dissimilatory. This is especially true in the case of high tones which may raise before low tones in Hokkien (Chang and Hsieh, 2012), Taiwanese (Peng, 1997), and Mandarin (Xu, 1997; Tilsen, 2013).
Why care about these patterns?

- If we wish to describe how tone is produced, our descriptions should be valid across contexts, including those where production is impinged upon.

- We know little about how context, rate, and stress interact in production, either with respect to segmental or tone production. Do patterns of tonal coarticulation mirror those of segmental coarticulation?

- Historically, tonal coarticulation is to tone sandhi what consonant coarticulation is to classical patterns of assimilation/dissimilation. Understanding tonal coarticulation may help explain how historical patterns of sandhi come about.
Method: stimuli

- How do tonal context and stress influence tone production in Triqui? How does rate? All three factors tested between words.

- Four tones were chosen and embedded in six different tonal contexts in natural sentences of 3 words.

- Example: \( nî^3 ?î^3 sîh^{45} ja^3 k^w eh^3 \) ‘The man knows Oaxaca.’

- The medial word was always monosyllabic while the adjacent words were always disyllabic.

- The resulting sentences were natural carrier sentences in the language of the form: Verb + Subject + Modifier (adj, object, etc.).
Experiment: tonal coarticulation

Method: tonal contexts and rate

Adjacent tone (L/R) | Medial tone
-------------------|-------------
/2.2/              | /45/        
/3.3/              | /4/         
/3.32/             | /43/        
/3.45/             | /2/         

Each of the 16 sentences were repeated five times by 8 speakers (4 male, 4 female) in two rate conditions (normal, fast).

Normal rate: 3.45 45 3.45, \texttt{[atJih\textasciitilde{}^{45} s\textasciitilde{}ih\textasciitilde{}^{45} t\textasciitilde{}a\textasciitilde{}^{3}kah\textasciitilde{}^{45}]}  
\textit{\textquoteleft El hombre pide el marrano.\textquoteright}  

Fast rate: 3.45 45 3.45, \texttt{[atJih\textasciitilde{}^{45} s\textasciitilde{}ih\textasciitilde{}^{45} t\textasciitilde{}a\textasciitilde{}^{3}kah\textasciitilde{}^{45}]}
The effect of contextual tones was coded by marking the difference in F$_0$ offset and onset between target vowels according to Chao's tone scale.

<table>
<thead>
<tr>
<th>Left</th>
<th>Medial</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>/2.2/</td>
<td>/45/</td>
<td>/2.2/</td>
</tr>
<tr>
<td>+2</td>
<td>-3</td>
<td></td>
</tr>
<tr>
<td>/3.45/</td>
<td>/2/</td>
<td>/3.45/</td>
</tr>
<tr>
<td>-3</td>
<td>+1</td>
<td></td>
</tr>
</tbody>
</table>
Measures

- Time-normalized $F_0$ data for the three vowels extracted using Voicesauce (Shue et al., 2009), at 20 points.

- $F_0$ data examined with several linear mixed effects model fit for each tone in each position (Left, Medial, Right) with 4 factors (Offset.difference X Time X Duration X Rate). Speaker was treated as a random effect with random intercepts and slope set for the effect of rate.

- The 20 time points were treated as continuous and recentered.

- Conservative measure of distance of z-score log $F_0$ from speaker’s average for a given tone (a measure of variation).

- All results discussed here are significant at $p < .01$ via model comparisons using a $\chi^2$ test with analysis of variance.
Guide to results

1. Tones on left syllable as a function of medial syllable tone.

2. Tones on medial syllable as a function of adjacent syllable tone.

3. Tones on right syllable as a function of medial syllable tone.

4. Rate effects

5. Duration differences
Left target: level tones

Anticipatory raising with T2, Left syllable

Time (normalized)

z-score normalized log F0

Following tone
- t2
- t4
- t43
- t45
Left target: contour tones

Assimilatory raising of T45, Left syllable
Summary - left target

- Significant effects of the following tone on the preceding tone’s F₀ trajectory, most notably for a following /45/ and /43/.

- Tone /2/ raises before contour tones /43, 45/, but is unaffected by following level tones.

- Tone /45/ is realized with an earlier F₀ peak before contour tones /43, 45/ than before following level tones.

- It is not simply the presence/absence of a higher tone which causes tonal coarticulation here, but the presence of a contour.
**Medial target: level tones**

Dissimilatory lowering of T2 and raising of T4

TIME (normalized)
**Medial target: contour tones**

![Graph showing dissimilatory raising of T45](image)

- Objective: Experiment on tonal coarticulation
- Medial target tones
- Data analysis: z-score normalized log F0 versus time (normalized)

**Adjacent tones**
- t2
- t3
- t32
- t45
Summary - medial target

- Significant effects of the adjacent tones on the medial tone’s F₀ trajectory, most notably for an adjacent /45/ and /32/.

- Not assimilatory, but dissimilatory.

- Tone /2/ lowers between contour tones /(3).45/ and tone /4/ raises between tones /2.2/

- Tone /45/ (and tone /43/ to a lesser extent) raises between tones /2.2/.
Right target: level tones

Assimilatory changes in tone /3/ in melodies /3.32/, /3.45/, Right syllable

Time (normalized)

Preceding tone

Following tone

Christian DiCanio (((Haskins)))

Triqui - coarticulation

4/6/14 29
Larger summary

- Strong anticipatory effects on word preceding medial monosyllable and strong carryover effects on following word too.

- Dissimilatory effects on word which varies in tone in the elicitation frame, but only with highest/lowest adjacent targets.

- Why dissimilate here?
Results II: Effect of rate on tone production

Average Normal: Fast speech rate = 1:1.26
Experiment: tonal coarticulation

Rate effect

**$F_0$ range expansion**

**Medial tones at different speech rates**

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2</td>
<td>T4</td>
<td>+T43</td>
<td>T45</td>
<td>Normal</td>
<td>Fast</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Christian DiCanio (((Haskins)))

Triqui - coarticulation

4/6/14 32
Two patterns are found in the data: a general increase in F₀ range during fast speech and a contraction during normal speech rate when a medial word falls between higher tones.

This range expansion is asymmetrical: higher tones are raised, but lower tones do not move.

<table>
<thead>
<tr>
<th>Left/Right tones</th>
<th>Rate</th>
<th>zF₀ range</th>
<th>Rate</th>
<th>zF₀ range</th>
</tr>
</thead>
<tbody>
<tr>
<td>/3.5/</td>
<td>normal</td>
<td>3.17</td>
<td>fast</td>
<td>4.77</td>
</tr>
<tr>
<td>/3.32/</td>
<td>normal</td>
<td>4.32</td>
<td>fast</td>
<td>4.36</td>
</tr>
<tr>
<td>/3.3/</td>
<td>normal</td>
<td>4.42</td>
<td>fast</td>
<td>4.63</td>
</tr>
<tr>
<td>/2.2/</td>
<td>normal</td>
<td>4.13</td>
<td>fast</td>
<td>4.73</td>
</tr>
</tbody>
</table>
**Duration differences for the examined vowels**

No difference in duration between stressed vowels, but unstressed vowels were shorter.

---

Christian DiCanio (((Haskins)))

Triqui - coarticulation

4/6/14 34
Tones did not differ substantially in duration in medial monosyllabic words.
Discussion: the effect of tonal context

- The strength of the coarticulatory effects was dependent on whether or not the following tone was contour/level.

- More dissimilatory effects were found in anticipatory contexts (medial), which is in-line with work on Taiwanese (Peng, 1997), Mandarin (Xu, 1998), Malaysian Hokkien (Chang and Hsieh, 2012), and Tianjin Chinese (Zhang and Liu, 2011).

- Dissimilation strongest with most divergent tone heights - tone /2/ and /45/ for instance. This is also the context where most dissimulation takes place with increased speech rate.
Tones underwent coarticulation, but they remained distinct from each other. This is different from Mandarin, where tones change drastically in context (Xu, 1994), but similar to findings in Thai (Gandour et al., 1999). Triqui and Thai both lack tone sandhi.

Suggests relationship between the presence of alternation and the degree of permitted variability in production (Gandour et al., 1999). Limited coarticulation preserves tonal contrasts.

Manuel’s output constraint: “languages generally tend to tolerate less contextually induced changes in acoustic phonetic output if they are likely to lead to confusion of contrastive phones.” (Manuel, 1990)
Discussion: Rate and contrast preservation

- Generally speaking, faster rate causes a shrinking of the $F_0$ range, as in French (Fougeron and Jun, 1998), English, and Croatian (Bradlow et al., 2003; Smiljanić and Bradlow, 2005).

- At a faster speech rate, mechanical constraints on $F_0$ production will produce greater assimilation and contrast neutralization.

- The findings here show range expansion during fast speech. Why?
Given the greater degree of coarticulation at faster speech rate, maintaining a pitch range may aid the listener in the perception of tonal contrasts, especially in a language without tone sandhi.

The preservation of contrasts requires the ability of a speaker to maintain stable $F_0$ trajectories across speech contexts. Triqui listeners are especially attuned to slope differences in tonal discrimination (DiCanio, 2012).

Languages for which increased variability in $F_0$ does not result in decreased lexical identification are not subject to the same constraint. The primary goals of speech perception are phonological and lexical categorization (Xu et al., 2006).
A general prosodic effect?

- Words under focus realized with $F_0$ range expansion in Mandarin Chinese (Liu and Xu, 2005; Xu, 1999) and Wenzhou (Scholz, 2012).

- “On contour tones, focus also affects the strength of the tonal realization, so that the tones are less influenced by adjacent tonal targets.” (Scholz, 2012, 189).

- General pattern occurs where tones are more resistant to contextual effects when in phonologically or pragmatically prominent positions.
Future directions

1. Modelling of assimilation and rate effects in general production model (TADA) under NSF grant (Whalen & Xu).

2. Investigate tonal coarticulation data from related language (Yoloxóchitl Mixtec).

3. Investigate tonal variability in Mixtec corpus data using forced alignment (DiCanio et al., 2012b, 2013).
Acknowledgements

The research was partly funded through a grant to Haskins Laboratories (Douglas Whalen, PI) on phonetic documentation in endangered languages.

Hosung Nam, Doug Whalen, and Caicai Zhang.

The Itunyoso Triqui community, kùruaa nihírèh! ([ku²ru⁴a⁴³ ni³ʔi⁴re?]!)

Thank you!


