The phonetics of prosody in Yoloxóchitl Mixtec

Christian DiCanio
cdicanio@buffalo.edu

Department of Linguistics
University at Buffalo

5/30/18
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
**An example**

t₃ ti¹ki¹⁴ yaa¹⁴ k₄⁴ ndi⁴ yaa¹⁴ sa⁴-ndu³ta³=ndu² t₄⁴ sa³k₄⁴ ndi⁴

Why do the instances of /yaa¹⁴/ ‘ash’ differ?
Prosody in endangered languages of Mexico

How does higher-level linguistic structure (information structure, intonation, boundaries) influence speech/tone production?

- Parallel speech production studies in the field with speakers of Itunyoso Triqui and Yoloxóchitl Mixtec
- Development of phonologically-annotated corpora in both languages

NSF Grant No. 1603323 to the University at Buffalo

Collaborators: Richard Hatcher, Basileo Martínez Cruz, Wilberto Martínez Cruz, Jonathan Amith, Rey Castillo García, Joshua Benn, Jason Lilley, Tim Bunnell
Roadmap

1. Background

2. Prosodic marking of focus (DiCanio et al., 2018a)

3. Boundary-adjacent lengthening and tonal effects (DiCanio et al. in prep)

4. Declination effects (DiCanio et al. in prep)

5. Discussion/Conclusion
How might focus be marked in a tone language?

1. **Phonological marking**
   Intonational pitch accents or boundary tones might overlap/influence tonal contour shapes.
   e.g. certain Swedish dialects (Bruce, 2005), Shekgalagari (Hyman and Monaka, 2011), Serbo-Croatian (Godjevac, 2005).

2. **Phonetic marking**
   Marked by phonetic lengthening, register shift, or pitch range expansion.
   e.g. Mandarin (Peng, 1997; Xu, 1999; Liu and Xu, 2005), Akan (Kügler and Genzel, 2011), Santa Ana del Valle Zapotec (Esposito, 2010).

3. **Only morphosyntactic marking**
   e.g. Northern Sotho (Zerbian, 2007), Itunyoso Triqui (DiCanio, in progress).
Register shift

High tones in Mandarin undergo raising and $F_0$ range expansion when in focus (Xu, 1999).

‘The kitty touches the kitty.’
Phonetic marking of domains

Domain-initial consonants may be lengthened or hyperarticulated (Fougeron and Keating, 1997; Keating et al., 2000)

The $F_0$ range may be expanded and articulatory gestures strengthened. (Mücke and Grice, 2014; Xu, 1999)

Stressed syllables may be the target of greater phonetic prominence (Byrd and Choi, 2010; de Jong, 1995; Krivokapić and Byrd, 2012)
Dynamical parameters (Cho, 2006)

(a) Change in Stiffness
(b) Change in Target
(c) Change in Intergestural Timing
(d) Change by shrinking

Figure 2. Hypothetical movement trajectories that correspond to a change in each parameter. (a) show change in stiffness; (b) change in target; (c) change in intergestural timing; and (d) change by shrinking. Empty circles indicate the timepoint of the peak velocity attainment.

1. Stiffness. Variation in articulatory movement duration is thought to be controlled by the stiffness parameter: the stiffer the spring (the articulator), the faster the movement. An idealized pattern in a pure change in stiffness is visualized in Figure 2a, and its corresponding kinematic relationships in Figures 3a-b. If stiffness is the only parameter underlying kinematic differences, there should be a change in peak velocity (i.e., the maximum velocity that the articulator attains during the gesture), but not in displacement (i.e., the amount of spatial distance that the articulator travels), therefore showing vertical distribution of the datapoints (Figure 3a). In addition, (with a change only in stiffness)
Research questions

1. How do tones change in prosodically weak/strong environments (e.g. under different focal conditions)?

2. How do tones change at prosodic boundaries?

3. Are there systematic phonetic differences between different types of prosodic effects (focus vs. boundary-adjacent lengthening)?
Yoloxóchitl Mixtec (YM)

- Otomanguean, spoken in Guerrero, Mexico (~4000 speakers).
- Phonological/phonetic fieldwork (Castillo García (2007), DiCanio et al. (2014, 2018a), DiCanio et al. (2018b), Palancar et al. (2016)).
Yoloxóchitl, Guerrero
All roots are minimally composed of bimoraic feet, consisting of either monosyllabic stems with long vowels (CVV) or disyllabic stems with shorter vowels (CVCV) (Castillo García, 2007).

No codas.

Glottalization is contrastive: /yaʔa¹/ ‘grey’, /saʔma⁴/ ‘cloth to wrap tortillas’

Final syllables are prominent.

- Nasal vowels only occur on stem-final syllables.
- Restricted vowel contrasts on non-final syllables.
- 9 tones on stem-final syllables, but only 5 on non-final syllables.
- Final syllable lengthening
Tone is lexical and morphological

Twenty-six tonal melodies are possible on a disyllabic word.

<table>
<thead>
<tr>
<th>Melody</th>
<th>Word</th>
<th>Gloss</th>
<th>Melody</th>
<th>Word</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>ta¹ma¹</td>
<td>without appetite</td>
<td>4.13</td>
<td>na⁴ma¹³</td>
<td>is changing</td>
</tr>
<tr>
<td>1.3</td>
<td>na¹ma³</td>
<td>to change (intr)</td>
<td>4.14</td>
<td>nda⁴ta¹⁴</td>
<td>is splitting up</td>
</tr>
<tr>
<td>1.4</td>
<td>na¹ma⁴</td>
<td>soap</td>
<td>4.24</td>
<td>ya⁴ma²⁴</td>
<td>Amuzgo person</td>
</tr>
<tr>
<td>1.32</td>
<td>na¹ma³²</td>
<td>I will change myself</td>
<td>4.42</td>
<td>na⁴ma⁴²</td>
<td>I often pile rocks</td>
</tr>
<tr>
<td>1.42</td>
<td>na¹ma⁴²</td>
<td>my soap</td>
<td>13.2</td>
<td>hi¹³ni²</td>
<td>has seen</td>
</tr>
<tr>
<td>3.2</td>
<td>na³ma²</td>
<td>wall</td>
<td>13.3</td>
<td>na¹³na³</td>
<td>has photographed (self)</td>
</tr>
<tr>
<td>3.3</td>
<td>na³ma³</td>
<td>to change (tr)</td>
<td>13.4</td>
<td>na¹³ma⁴</td>
<td>has piled rocks</td>
</tr>
<tr>
<td>3.4</td>
<td>na³ma⁴</td>
<td>sprout</td>
<td>14.2</td>
<td>na¹⁴ma²</td>
<td>I will not change</td>
</tr>
<tr>
<td>3.42</td>
<td>na³ma⁴²</td>
<td>I will pile rocks</td>
<td>14.3</td>
<td>na¹⁴ma³</td>
<td>to not change</td>
</tr>
<tr>
<td>4.1</td>
<td>ka⁴nda¹</td>
<td>is moving (intr)</td>
<td>14.4</td>
<td>na¹⁴ma⁴</td>
<td>to not pile rocks</td>
</tr>
<tr>
<td>4.2</td>
<td>na⁴ma²</td>
<td>I am changing</td>
<td>14.13</td>
<td>na¹⁴ma¹³</td>
<td>to not change oneself</td>
</tr>
<tr>
<td>4.3</td>
<td>na⁴ma³</td>
<td>it is changing</td>
<td>14.14</td>
<td>nda¹⁴ta¹⁴</td>
<td>to not split up</td>
</tr>
<tr>
<td>4.4</td>
<td>na⁴ma⁴</td>
<td>is piling rocks</td>
<td>14.42</td>
<td>na¹⁴ma⁴²</td>
<td>I will not pile rocks</td>
</tr>
</tbody>
</table>
## Morphological tone

<table>
<thead>
<tr>
<th>Morphology</th>
<th>‘to break’ (tr)</th>
<th>‘hang’ (tr)</th>
<th>‘to change’ (intr)</th>
<th>‘to peel’ (tr)</th>
<th>‘to get wet’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stem</td>
<td>tə³ʔi⁴</td>
<td>tʃi³kū²</td>
<td>nɑ¹ma³</td>
<td>kwi¹i⁴</td>
<td>tʃi³i³</td>
</tr>
<tr>
<td>NEG</td>
<td>tə¹⁴ʔi⁴</td>
<td>tʃi¹⁴kū²</td>
<td>nɑ¹⁴ma³</td>
<td>kwi¹⁴i¹⁴</td>
<td>tʃi¹⁴i³</td>
</tr>
<tr>
<td>COMP</td>
<td>tə¹³ʔi⁴</td>
<td>tʃi¹³kū²</td>
<td>nɑ¹³ma³</td>
<td>kwi¹⁴i⁴</td>
<td>tʃi¹³i³</td>
</tr>
<tr>
<td>INCOMP</td>
<td>tə⁴ʔi⁴</td>
<td>tʃi⁴kū²</td>
<td>nɑ⁴ma¹³</td>
<td>kwi⁴i¹⁴</td>
<td>tʃi⁴i⁴</td>
</tr>
<tr>
<td>1S</td>
<td>tə³ʔi⁴²</td>
<td>tʃi³kū²=ju¹</td>
<td>nɑ¹ma³²</td>
<td>kwi¹⁴²</td>
<td>tʃi³i²</td>
</tr>
</tbody>
</table>
How do we elicit information structure in YM?

- Illiterate population, so a reading task will not work. c.f. studies on Mandarin (Chen and Gussenhoven, 2008; Xu, 1999), Guaraní (Clopper and Tonhauser, 2013), Arabic (de Jong and Zawaydeh, 2002), German (Mücke and Grice, 2014), or Dutch (Peters et al., 2014).

- Mining a corpus for examples does not control for tone or word structure.

- A Q&A paradigm following a short story elicits NPs with different information structure, but this does not work well for broad focus. c.f. studies on Akan (Kügler and Genzel, 2011), Guaraní (Clopper and Tonhauser, 2013),
Stimuli elicitation for focus - a mixed design

- **Argument focus (after story)**
  Rey: Who arrived?
  Speaker: John arrived.

- **Contrastive focus (after story)**
  Rey: Did Marcus arrive?
  Speaker: John arrived.

- **Sentential focus (repetition)**
  Rey: John arrived.
  Speaker: John arrived.
Focus in Yoloxóchitl Mixtec

(1) ni₁-ta³ji³ yu³βa⁴=ð⁴ kwa⁴yu² nda³?a⁴=ð⁴
PERF-give father=2S horse hand=2S
‘Your father gave you a horse.’

(2) yu³βa⁴=ð⁴ ni¹-ta³ji³=ri⁴ kwa⁴yu² nda³?a⁴=ð⁴
father=2S PERF-give=3S horse hand=2S
‘Your father gave you a horse.’

(3) yu³βa⁴=ð⁴ ni¹-ta³ji³=ri⁴ kwa⁴yu² nda³?a⁴=ð⁴
father=2S PERF-give=3S horse hand=2S
‘Your father gave you a horse.’
Experiment 1: Focus and stress

Methods

- Each answer/response was repeated six times by each respondent across two separate recording sessions (3 reps/session).
- Recording took place in San Luis Acatlán, a town near Yoloxóchitl.
- Each condition contained the same 28 target words which possessed nine tonal melodies: 1.1, 1.3, 1.4, 1.42, 3.2, 3.3, 3.4, 4.2, 4.4; all disyllables.
- Ten native speakers participated; a total of 2,595 utterances were analyzed.
- Target words segmented and analyzed using a script written in Praat (Boersma and Weenink, 2016).
- Normalized F0 trajectories extracted over 5 time points and converted to log-normal values. Onset and vowel duration also extracted.
- Results analyzed using LMMs with lmertest (Kuznetsova et al., 2017). All reported results are significant.
Since we have a mixed design, we will present the contrastive-argument focus comparison first and then compare them both to the sentential focus condition.
Results: Duration I

Onset duration by stress and focus type

Vowel duration by stress and focus type

Word position
unstressed
stressed

Focus
argument
contrastive

Focus
argument
contrastive
Results: Duration - comparative

Onset duration by stress and focus type

Vowel duration by stress and focus type

Word position
- unstressed
- stressed
### Experiment 1: Focus and stress

#### Results

<table>
<thead>
<tr>
<th></th>
<th>$C_1$</th>
<th>$V_1$</th>
<th>$C_2$</th>
<th>$V_2$</th>
<th>$\sigma_1$</th>
<th>$\sigma_2$</th>
<th>$\sigma$-duration ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline (sentential)</strong></td>
<td>70</td>
<td>77</td>
<td>95</td>
<td>90</td>
<td>141</td>
<td>185</td>
<td>1:1.31</td>
</tr>
<tr>
<td><strong>Contrastive focus</strong></td>
<td>77</td>
<td>92</td>
<td>120</td>
<td>99</td>
<td>169</td>
<td>219</td>
<td>1:1.30</td>
</tr>
<tr>
<td>comparison to baseline</td>
<td>10%</td>
<td>19%</td>
<td><strong>26%</strong></td>
<td>10%</td>
<td>20%</td>
<td>18%</td>
<td></td>
</tr>
<tr>
<td><strong>Argument focus</strong></td>
<td>76</td>
<td>94</td>
<td>136</td>
<td>107</td>
<td>170</td>
<td>242</td>
<td>1:1.42</td>
</tr>
<tr>
<td>comparison to baseline</td>
<td>9%</td>
<td>22%</td>
<td><strong>43%</strong></td>
<td>19%</td>
<td>21%</td>
<td>31%</td>
<td></td>
</tr>
</tbody>
</table>

Final syllables are longer than penults. Under focus, greater lengthening occurs in the onset of the stressed syllable than in the vowel.
Results: Level and rising melodies

Globally, contrastive focus undergoes raising relative to argument focus.

**Effect of focus type on level tonal melodies**

<table>
<thead>
<tr>
<th>Focus</th>
<th>non-final</th>
<th>final</th>
</tr>
</thead>
<tbody>
<tr>
<td>argument</td>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
</tr>
<tr>
<td>contrastive</td>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
</tr>
</tbody>
</table>

**Effect of focus type on rising tonal melodies**

<table>
<thead>
<tr>
<th>Focus</th>
<th>non-final</th>
<th>final</th>
</tr>
</thead>
<tbody>
<tr>
<td>argument</td>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
</tr>
<tr>
<td>contrastive</td>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
</tr>
</tbody>
</table>

DiCanio et al (UB) Prosody in YM 5/30/18 24 / 52
Results: Tone - comparative

For most melodies, tone is lower in sentential focus than argument focus.

Effect of focus type on level tonal melodies

Effect of focus type on rising tonal melodies
Discussion: duration and focus

Focus lengthens the onset of the stressed syllable more than the vowel. Why?

- In Swedish onsets are lengthened when a syllable contains a phonologically short vowel (Heldner and Strangert, 2001).

- Vowels in the CVCV disyllables were short, so vowel length may have influenced the domain of prosodic lengthening in YM.
Discussion: tone and focus

- Tones in fronted, focal NPs undergo F0 range expansion and raising relative to tones in sentential focus.

- Contrastive focus has the most raising.

- Tone /1/ is specifically not raised when it is the initial tone in a rising sequence on a disyllable, e.g. 1.4, 1.42, 1.3.

- Focus induces processes of tonal hyperarticulation that enhance syntagmatic contrast on the word.
Asymmetrical expansion occurs because low tones are near the F₀ floor (c.f. high vowel displacement under different focus conditions (Cho, 2006; Mücke and Grice, 2014)).
Phrase-final phenomena

How are tones in YM influenced by phrasal position?

Phrase-final position is a domain of articulatory strengthening and where articulatory gestures may reduce their velocity. (Barnes, 2006; Cho, 2006; Krivokapić and Byrd, 2012).

Declination is a universal phonetic process but phrase-final tonal alternations may be phonological (Gussenhoven, 2004).

Can we separate domain-final effects from global effects in speech production?
Is it actually just a domain-final effect?

Pike and Small (1974); Pike and Wistrand (1974) provide only impressionistic comments regarding positional differences.
Declination, final lowering, and tone languages

Declination is a universal process in declarative utterances (Gussenhoven, 2004), but there are exceptions in tone languages:

1. It does not occur in a sequence of high tones, e.g. Mandarin (Xu, 1999), Taiwanese (Peng, 1997).
2. It only occurs in a sequence of low tones, e.g. Mambila (Connell, 2017), Yoruba (Laniran and Clements, 2003).
3. It does not occur, e.g. Choguita Rarámuri (Garellek et al., 2015), Embosi (Rialland and Embanga Aborobongui, 2017).

Final lowering occurs in tone languages:

1. It occurs for all tones, e.g. Kipare (Herman, 1996), Moro (Chung et al., 2016), Embosi (Rialland and Embanga Aborobongui, 2017)
2. It only occurs with low and falling tones, e.g. Mambila (Connell, 2017), Taiwanese (Peng, 1997), Akan (Kügler, 2017).
Methods: positional effects on tone

- 20 tonal melodies were analyzed (1.1, 1.3, 1.42...) in disyllabic words in non-final contexts (before a PP/Adv) and utterance-final contexts.

\[ Sa^4Si^{24} = Ra^2 ni^3Si^4 \] 'He is eating corn.'

\[ Sa^4Si^{24} = Ra^2 ni^3Si^4 \beta i^3ti^3 \] 'He is eating corn now.'

- The post-target word always had tone /3/.
- 288 repetitions for each speaker (36 words x 2 conditions x 4 repetitions); 9 speakers.
- Initial transcription in ELAN and segmentation in Praat. We used a script to analyze $F_0$ dynamics and duration.
- $F_0$ was normalized and all data was analyzed using the same methods as experiment 1.
Results II: duration

Onset duration by stress and utterance position

Vowel duration by stress and utterance position
Results II: level tone melodies

Effect of sentence position on level tonal melodies /1.1, 3.3, 4.4/

- Higher
- Lower
Results II: falling and rising melodies

Effect of sentence position on tonal melodies

<table>
<thead>
<tr>
<th>Penult</th>
<th>Ultima</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="#" alt="Graph for falling melodies" /></td>
<td><img src="#" alt="Graph for rising melodies" /></td>
</tr>
</tbody>
</table>

Utterance position
- non-final
- final

Tonal melody
- 3.2
- 4.1
- 4.2
- 4.3

Final raising with /1.4, 3.4/

Final lowering with /3.2, 4.1/
Results II: melodies with final contours

Effect of sentence position on tonal melodies /1.42, 3.42, 1.32/

Effect of sentence position on tonal melodies /4.24, 4.13/

- **Final lowering of /42, 32/**
- **Levelling allotony:** /24/ > 3/4_T
- **Utterance position**
  - non-final
  - final
- **Tonal melody**
  - 1.42
  - 3.42
  - 1.32
- **Allotony:** /13/ > 42/4_T

DiCanio et al (UB)
Discussion

1. Vowels are lengthened in phrase-final position.

2. Tonal effects occur only in the boundary-adjacent syllable.

3. Phrase-final position is marked by F₀ range expansion. The highest tone /4/ raises and lower/falling tones (/2, 1, 42, 32/) lower. Tone /3/ does not change.

4. Rising tones (/13, 24/) have distinct allotones in non-utterance-final position.

Are processes in final position related to utterance-level declination or raising?
Methods: Declination

- We analyzed sentences that consisted of only level tone sequences.
- 10 sentences between 4-7 syllables in length; 2 with tone /4/, 4 with tone /3/, 4 with tone /1/.
- 10 sentences × 4 repetitions; 9 speakers.
- Initial transcription in ELAN and segmentation in Praat. We used a script to analyze F₀ dynamics and duration.
- F₀ was normalized and all data was analyzed using the same methods as experiment 1.
- Two statistical methods to disambiguate declination from final lowering: trajectory modelling with and without utterance-final syllable.
Results - declination

Occurs in sequences of tone /1/ and /3/, but not with tone /4/.

F0 change across utterances with identical level tones, medial 60% of vowel
Results - declination modelling

F0 change across utterances with identical level tones, medial 60% of vowel.
Solid lines = entire lmer fit; Black, dashed lines = fit without final syllable

Normalized log(F0)

Syllable number from end of utterance

Tone

/4/

/3/

/1/
Processes affecting final tones

- Final raising of highest tone and lowering of lowest tone reflect distinct processes from utterance-level effects.

- Utterance-level declination occurs with non-high tones but not with the highest tone (/4/).

- Are these boundary tones? No. If they were to exist, we would have to stipulate that they be extensions of the same preceding tones, i.e. H% only after /4/. 

![Diagram showing non-final and final position changes in tones]
Conclusions: multiple prosodic mechanisms

The type of F₀ range expansion and durational changes observed on initial focused constituents are distinct from those observed in phrase-final position.

Prosodic marking of focus in YM is distinct from boundary-related prosodic effects.

Tonal changes in utterance-final position result from tonal hyperarticulation which expands the tonal range (Krivokapić and Byrd, 2012).
Duration and $F_0$ raising on tone /4/ is correlated, but it is not a strong relationship.

Relationship between vowel duration and $F_0$ maximum in word-final syllables
$R=0.24$
Conclusions: mechanisms

Prosody in YM is marked primarily by adjustments to $F_0$ range and hyper/hypoarticulation (de Jong, 1995; de Jong and Zawaydeh, 2002).

(b) Change in Target

Change in range = postural target adjustment?

Onset

H
M
L

H
M
L

larger Target

gesture 1

gesture 2

smaller Target

H

M

L
Future plans

1. Parallel research on Itunyoso Triqui (IT) prosody.

2. Tone production in the YM and IT corpora.

3. EMA research in the UB Phonlab on the supralaryngeal articulation of information structure in English and Korean.
Acknowledgements

- Support from NSF DEL/RI grant 1603323, *Understanding Prosody and Tone Interactions through Documentation of Two Endangered Languages*

- Team Mixtec: Rey Castillo García (SEP, México), Jonathan Amith (Gettysburg College), and Joshua Benn (University at Buffalo)

- Commentary from audiences at CILLA VIII and UC Santa Cruz.
Results: Falling and complex melodies

Effect of focus type on falling tonal melodies

Effect of focus type on complex tonal melody /1.42/
Results - declination by speaker

F0 change across utterances with identical level tones, medial 60% of vowel; by speaker
Variation in the production of rising tones

Effect of sentence position on tonal melodies /4.24, 4.13/, by speaker

Speakers who maintain /24/ as a rising tone.

Utterance position
- non-final
- final

Tonal melody
- 4.24
- 4.13

Speakers who maintain both rising tones.
Final allotony

$F_0$ rises require more time than level or falling trajectories, thus we might expect that they be limited to contexts with longer phonetic duration, e.g. phrase-final position (Sundberg, 1979; Zhang, 2004).

Allotony results from durationally-induced $F_0$ levelling. Levelling is induced via articulatory undershoot (Parrell, 2014; Mücke and Grice, 2014).
Prosodic marking

**Accentual marking of heads/edges** – intonational pitch accents are attracted to prominent positions in the prosodic hierarchy or on constituents with narrow focus (Gussenhoven, 2004; Pierrehumbert and Beckman, 1988).

**Non-accentual phonological marking of domains** – prominent positions in the prosodic hierarchy license the application of specific phonological processes, e.g. tone spreading domains (Hsu and Jun, 1996; Hyman, 1990; Hyman and Monaka, 2011; Lee, 2014), positional neutralization (Barnes, 2006).

**Phonetic marking of domains** – prominent positions in the prosodic hierarchy undergo processes of phonetic enhancement, e.g. domain-initial strengthening (Fougeron and Keating, 1997; Keating et al., 2000), focal $F_0$ range expansion (Xu, 1999), stress-related hyperarticulation (Byrd and Choi, 2010; de Jong, 1995; Krivokapić and Byrd, 2012).
Duration and F₀ lowering on tone /1/ are negatively correlated, but this is a weak effect.


