Is intonation universal?

Christian DiCanio
cdicanio@buffalo.edu
Richard Hatcher
rjhatche@buffalo.edu

Department of Linguistics
University at Buffalo

11/30/18
Biological bases for pitch variation

Pitch variation in speech is determined by three universal, biological codes (Gussenhoven, 2004, 2016):

- **Frequency code**: pitch height is associated with dominance (Ohala, 1983, 1994)
- **Effort code**: articulatory effort ($F_0$ height) is iconic with degree of linguistic emphasis
- **Production/respiratory code**: Boundary-related $F_0$ effects are associated with breathing and are related to topicality, finality, etc.
How languages implement the effort and production codes is language-specific, but the codes are biological and, ostensibly, universal.

- Pitch accents occur on words marked with narrow focus in many languages (English, German, Italian, Swedish, etc). This is a grammaticalization of the effort code.

- Boundary tones (e.g. L%, H%) are language-specific implementations of the production code.

Though some tonal languages show evidence of the effort code, the motivation for these biological codes is based on work on non-tonal languages and there are counter-examples (in tone languages) (c.f. Kügler and Genzel (2011)).
Introduction

Is there intonation in tone languages?

- Pitch accents are either minimal or do not occur.
  e.g. Mandarin (Xu, 1997), Mambila (Connell, 2017), Yoloxóchitl Mixtec (DiCanio et al., 2018), Yoruba (Laniran and Clements, 2003)

- Boundary tones may be absent or may only co-occur with certain tones.
  e.g. Akan (Kügler, 2017), Bàsàá (Makasso et al., 2017), Mandarin (Xu, 1999), Taiwanese (Peng, 1992), Tswana (Zerbian, 2017), Yoruba (Laniran and Clements, 2003)

→ Intonational effects may be phonetically layered on existing lexical tones and cause (a) $F_0$ register shift or (b) $F_0$ range fluctuation.
  e.g. Mandarin (Xu, 1999), Yoloxóchitl Mixtec (DiCanio et al., 2018)
High tones in Mandarin undergo raising and F₀ range expansion when in focus (Xu, 1999).

‘The kitty touches the kitty.’
Range expansion in utterance-final position

Expanded $F_0$ range of Mixtec tones (DiCanio et al, in progress).

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

- Higher
- Lower
Itunyoso Triqui (Otomanguean) possesses a complex tonal system and does not possess either pitch accents or boundary tones.

Does the language show evidence for these intonational effects elsewhere, such as in the realization of narrow/contrastive focus and at utterance-boundaries?
The Triqui region
Most morphemes (73% of roots) are polysyllabic.

Nine lexical tones contrast on final syllables. Tone in non-final syllables is often redundant (e.g. \[ru^4ne^43\] ‘avocado’) but may be contrastive (/2/ vs. /3/, /3/ vs. /4/) (DiCanio, 2008, 2016).

<table>
<thead>
<tr>
<th>Tone</th>
<th>Open syllable</th>
<th>Coda /h/</th>
<th>Coda /ʔ/</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Word</td>
<td>Gloss</td>
<td>Word</td>
</tr>
<tr>
<td>/4/</td>
<td>yū⁴</td>
<td>‘earthquake’</td>
<td>yāh⁴</td>
</tr>
<tr>
<td>/3/</td>
<td>yū³</td>
<td>‘palm leaf’</td>
<td>yāh³</td>
</tr>
<tr>
<td>/2/</td>
<td>ū²</td>
<td>‘nine’</td>
<td>tah²</td>
</tr>
<tr>
<td>/1/</td>
<td>yū¹</td>
<td>‘loose’</td>
<td>kāh¹</td>
</tr>
<tr>
<td>/45/</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/13/</td>
<td>yo¹³</td>
<td>‘fast (adj.)’</td>
<td>toh¹³</td>
</tr>
<tr>
<td>/43/</td>
<td>rā⁴³</td>
<td>‘want’</td>
<td>nnāh⁴³</td>
</tr>
<tr>
<td>/32/</td>
<td>rā³²</td>
<td>‘durable’</td>
<td></td>
</tr>
<tr>
<td>/31/</td>
<td>rā³¹</td>
<td>‘lightning’</td>
<td></td>
</tr>
</tbody>
</table>
## Triqui grammar/phonology

- Final syllables are bimoraic, consisting of the shapes /CVh, CV?, CV:/, and prominent. Most of the phonological contrasts occur on them (DiCanio, 2008).
- Tone has a high morphological load in the language, marking person, verbal aspect, and a few other distinctions (DiCanio, 2016).

<table>
<thead>
<tr>
<th>Triqui Morphology</th>
<th>English Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>tʃa⁴³</td>
<td>'to eat (PERF)'</td>
</tr>
<tr>
<td>tʃah⁴</td>
<td>'I ate'</td>
</tr>
<tr>
<td>tʃa⁴¹ = reʔ¹</td>
<td>'You ate'</td>
</tr>
<tr>
<td>tʃah³</td>
<td>'(aforementioned) ate'</td>
</tr>
<tr>
<td>tʃoʔ⁴</td>
<td>'We ate'</td>
</tr>
<tr>
<td>tʃa²</td>
<td>'to eat (POT)'</td>
</tr>
<tr>
<td>tʃah¹</td>
<td>'I will eat'</td>
</tr>
<tr>
<td>tʃah²³</td>
<td>'(aforementioned) will eat'</td>
</tr>
<tr>
<td>tʃoʔ²</td>
<td>'We will eat'</td>
</tr>
</tbody>
</table>
All words are tonally marked in Triqui and there are no pitch accents.

<table>
<thead>
<tr>
<th>ka³</th>
<th>yu³</th>
<th>ne³</th>
<th>tā³</th>
<th>ri³</th>
<th>ki³</th>
<th>tji³</th>
<th>lu³</th>
</tr>
</thead>
<tbody>
<tr>
<td>ka³yu³</td>
<td>ne³tā³</td>
<td>ri³ki³</td>
<td>tji³lu³</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

fell | beans | beneath | cat

*The beans fell under the cat.*
Much of the pragmatic work usually done by intonation is handled by obligatory utterance-final particles (at least 24 of them). These do not seem to influence the $F_0$ of preceding tones.

‘Are you going to buy some huipil pieces?’ (clipped)

‘You are going to buy some huipil pieces apparently.’ (clipped)
Open questions

1. Might tones have alternate realizations under different information structure contexts? (effort code)

2. Are there boundary tones? What happens to tones at utterance boundaries? (production code)

Is there any evidence for intonation?
What might we expect?

- **Focus** may be 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).

- **Final lowering** may occur for all tones or be restricted to low/falling tones.
  All tones: Kipare (Herman, 1996), Moro (Chung et al., 2016), Embosi (Rialland and Embanga Aborobongui, 2017)
  Low/falling tones: Mambila (Connell, 2017), Taiwanese (Peng, 1997).

- **Declination** is limited to a sequence of high or low tones; or be absent.
  Restricted: Mandarin (Xu, 1999), Taiwanese (Peng, 1997), Mambila (Connell, 2017), Yoruba (Laniran and Clements, 2003)
  Absent: Choguita Rarámuri (Garellek et al., 2015), Embosi (Rialland and Embanga Aborobongui, 2017).
Eliciting information structure in Itunyoso Triqui

- 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)).

- A mixed design was used; both repetition and a Q&A paradigm (c.f. (DiCanio et al., 2018)).
Why a mixed design?

1. Itunyoso Triqui uses pronominal clitics for animate entities that have been backgrounded.

2. Mixtecan languages are object-dropping, so mentioned objects are absent.

3. “What happened?” is an odd question after listening to a text. Speakers attempt to answer it by speculating about the actor’s intentions in the story.
Methods

- Each answer/response was repeated five times by each speaker; 3 conditions (broad focus, contrastive focus, narrow focus).
- Recording took place in Tlaxiaco, Mexico and San Martín Itunyoso.
- Each condition contained the same 50 target words which possessed tones /1, 2, 3, 4, 45, 13, 32, 43/ on monosyllables and disyllables, with each rime type (/V:, Vf, V?/).
- 11 native speakers participated; a total of 8250 utterances were analyzed.
- Target words segmented and analyzed using a script written in Praat (Boersma and Weenink, 2016).
- Normalized F$_0$ trajectories extracted over 5 time points and converted to log-normal values. Syllable duration also extracted.
- Results analyzed using LMMs with lmertest (Kuznetsova et al., 2017).
Results: Duration

Syllables are longer under narrow/contrastive focus than under broad focus.

Syllable duration by coda type and focus condition in monosyllables

Syllable duration by coda type and focus condition in disyllables

DiCanio et al (UB)

Triqui intonation

11/30/18 18 / 40
Results: Tone in monosyllabic words

Effect of focus type on level tones

Effect of focus type on rising tones

Effect of focus type on falling tones

DiCanio et al. (UB)
Results: Tone in glottal contexts

Effect of focus type on level tones in monosyllabic words in different rime contexts

Tonal raising

Time (normalized)

DiCanio et al (UB)  Triqui intonation
Interim discussion - no effect of focus on $F_0$

No general effect of narrow/contrastive focus on $F_0$ across tonal categories, but a significant effect for tones with a coda /ʔ/.

Words are longer under contrastive and narrow focus than under broad focus; open syllables lengthen more (20%) than Vʔ and Vʰ syllables (5-15%).

Tone-specific effects of information structure occurred (tone /4/, /3/), but of relatively small magnitude (0.25 - 0.5 s.d.)
F₀ is lower under broad focus for /Vʔ/ rimes. Why?

Coda /ʔ/ induces F₀ lowering on tone (c.f. DiCanio (2012a)) and these effects might be weaker under narrow or contrastive focus; where speech is hyperarticulated. See figure from DiCanio (2012a, 170):

![Figure 9: F₀ values for tones in /V:/ and /Vʔ/ conditions.](image-url)
Focus involves the avoidance of glottalization-induced $F_0$ perturbations; tones are uttered more carefully.

There is no general effect of focus on $F_0$ of tones, but perhaps “effort” in IT is implemented via lengthening and other cues.

What about boundary-related effects?
10 tonal melodies were analyzed (3.5, 4.4, 4.43...) in disyllabic words in non-final contexts (before a PP/Adv) and utterance-final contexts.

\[
\begin{align*}
\text{ki}^3 \text{rāh}^4 & \text{ neh}^3 \text{ ṭjī}^3 \text{nāh}^5 \text{ ‘They bought a huipil.’} \\
\text{ki}^3 \text{rāh}^4 & \text{ neh}^3 \text{ ṭjī}^3 \text{nāh}^5 \text{ ni}^3 \text{γjāh}^5 \text{ ‘They bought a huipil in Tlaxiaco.’} \\
\end{align*}
\]

The pre-target word always had tone /2/. The post-target word always had tone /3/.

400 repetitions for each speaker (50 words x 2 conditions x 4 repetitions); 10 speakers (5M/5F)

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.
**Duration**

Syllable duration by word and utterance position

![Box plot showing syllable duration](image)

- **Non-final position**: 500 ms
- **Final position**: 300 ms

- **Penult** vs. **Final**
- **/r/** vs. **/h/** vs. **No coda**

**Results**

DiCanio et al (UB)

Triqui intonation

11/30/18 25 / 40
Experiment 2: Phrasal position and tone

Results

Results - tones in disyllabic words

Effect of sentence position on tonal melodies /3.3, 2, 2, 1.1/

Effect of sentence position on tonal melodies /1.3, 3.4, 3.5/

Effect of sentence position on tonal melodies /3.1, 3.32, 4.43/

Slope and F0 level are significantly different across utterance positions.
Summary - little effect of position on $F_0$

Final syllables are longer than non-final syllables and lengthened in utterance-final position.

As in the focus data, open syllables were lengthened more than closed syllables (1:1.5 vs. 1:1.37, 1:1.12).

Minimal effect of utterance position on $F_0$ of tones /4.4, 3.32, 3.4/. No effect on any other tonal melody.

However, investigating the slope on the falling tones across utterance positions revealed them to be equivalent.
Discussion: where is intonation in Itunyoso Triqui?

The $F_0$ of tones is unaffected by changes to information structure or utterance position.

Prosody influences syllable duration and this may, in turn, permit speakers a larger durational window for the hyperarticulation of contrasts on the word (c.f. DiCanio et al. (2018) on Yoloxóchitl Mixtec).

In Itunyoso Triqui, $F_0$ does not appear to be one of the parameters which is hyperarticulated in the examined contexts.

Speakers may be inconsistent in their use of pitch accents (Grice et al., 2017) but consistent in supralaryngeal hyperarticulation (Mücke and Grice, 2014).
**Functional load of $F_0$ and duration?**

$F_0$ varies not only with the dense lexical tone system, but also with coda glottal consonants (DiCanio, 2008, 2012a).

Prosodic lengthening is restricted since length is phonemic in consonants (DiCanio, 2012b), a strong cue to coda glottal consonants (DiCanio, 2014), and varies with tone (DiCanio, 2008).

Word-prosodic complexity restricts the degrees of freedom for the phonetic realization of intonation in Itunyoso Triqui.
But functional load of tone must not be the whole story. In Yoloxóchitl Mixtec, a related language with a dense tonal inventory, we observe a significant effect of focus on tonal range (DiCanio et al., 2018).

And we used identical methods!
And there are also positional effects in Yoloxóchitl Mixtec, unlike in Itunyoso Triqui. High tones raise; low and falling tones lower.

(DiCanio et al, in progress) - again, we used the same methods!
Cross-linguistic comparison

Both languages have 4 tonal levels, contour tones, and tonal morphology involving aspect/person.

Yet, Itunyoso Triqui possesses a few additional, grammatical mechanisms that encode things otherwise marked intonationally. There are many final particles with pragmatic functions (Mixtec has fewer):

<table>
<thead>
<tr>
<th>Type</th>
<th>Particle</th>
<th>Type</th>
<th>Particle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interrogative</td>
<td>niʔ⁴, oʔ¹, ah³, ūh⁴, aʔ³, saʔ¹</td>
<td>Evidential</td>
<td>rah¹, nãʔ³, reh³, sa³ yoh³</td>
</tr>
<tr>
<td>Command/finality</td>
<td>nāh⁵, βeh¹, sah⁵</td>
<td>Declarative</td>
<td>yu³ βe³², na² yoh³</td>
</tr>
<tr>
<td>Directive</td>
<td>?neh⁵, ya³ rih⁵</td>
<td>Negative</td>
<td>ya³ meh³, māh⁵</td>
</tr>
</tbody>
</table>

DiCanio et al (UB)
Morphologized focus/topic

\[ \text{ra}^3\text{ʔa}^3 \rightarrow \text{ra}^3\text{ʔah}^3 \quad \text{nāh}^{45} \rightarrow \text{nā}^3 \]

'hand' \quad 'hand.3TOP' \quad 'wash' \quad 'wash.3TOP'

Topical 3\textsuperscript{rd} person: toggle \( \varnothing \leftrightarrow \text{h} \) on root and replace final tone with /3/.

\[ \text{tʃa}^3\text{ʔo}^3 \rightarrow \text{tʃa}^4\text{ʔo}^4 \quad \text{tʃa}^4^3 \rightarrow \text{tʃa}^4^4 \]

'rabbit' \quad 'Rabbit, that rabbit' \quad 'eat.PERF' \quad 'THEY ate'

Emphatic tone: add -/ʔ/ on root and replace final tone with /4/.
-used for disjoint reference, contrastive topic, disambiguation
The presence of these additional mechanisms in Itunyoso Triqui for marking topicality and disambiguating referents may make it qualitatively different from Yoloxóchitl Mixtec.

Perhaps it is not the number of tonal contrasts which determines the presence of intonational effects, but the number of additional mechanisms available for highlighting or de-emphasizing constituents.

But this is speculative - more work is needed!
Conclusions

- $F_0$ is not used for marking focus or utterance-final boundaries in Itunyoso Triqui.

- Insofar as the **effort code** and the **production code** are universals grounded in biology, then the data here is a counter-example.

- Yet, perhaps such codes are too specifically tied to $F_0$; prosody equally involves lengthening and hyperarticulation.

- Is the grammaticalization of pragmatic meaning more common in complex tonal languages?
Future plans

1. Inclusion of other factors that speakers may be manipulating (spectral tilt, intensity).

2. Research on declination in utterances with varying final particles.

3. Corpus tone production in parallel annotated corpora of Yoloxóchitl Mixtec and Itunyoso Triqui.

4. 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 Triqui: Basileo Martínez Cruz, Wilberto Martínez Cruz, the Itunyoso Triqui community

[NSF National Science Foundation logo]
Stimuli elicitation for focus - a mixed design

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

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

- Sentential focus (repetition)
  Consultant: John arrived.
  Speaker: John arrived.
Appendix

Results: Tone in disyllabic words - focus

Effect of focus type on level melodies in disyllables

Effect of focus type on rising melodies in disyllables

Effect of focus type on falling melodies in disyllables

DiCanio et al (UB)  Triqui intonation  11/30/18  39 / 40
Results: Positional effects by coda type

Effect of sentence position on tonal melodies /4.4, 3.3, 2.2, 1.1/ by final coda type

- /ʔ/: penult, final
- /h/: penult, final
- no coda: penult, final

Z-score normalized log(F0)

Utterance position
- non-final
- final

Tonal melody
- T4.4
- T3.3
- T2.2
- T1.1

Time (normalized)


References


References


