Assessing intonational grammars through simulation and classification of pitch trajectories: the case of mobile boundary tones in Blackfoot

Natalie Weber (natalie.weber@yale.edu)
Jason Shaw (jason.shaw@yale.edu)

Yale University

June 25, 2022
A standard LabPhon method

- Evaluate theory through **experiments**
  - isolate variables of interest in two or more conditions
  - control other factors

- Problems? Not every question is well-suited to these methods.
  - assumes some uniformity of items within a condition
  - difficult in intonational phonology, as there tends to be a many-to-many mapping between semantic categories and pitch accents (Im, Cole & Baumann 2018, Roettger, Mahrt & Cole 2019)
Alternative method: simulation and classification

- Alternative: evaluate competing theoretical proposals via *simulation and classification* (Shaw & Kawahara 2018, Kawahara, Shaw & Ishihara 2021)
  - create stochastic generative models for competing theoretical hypotheses
  - use models to assign posterior probabilities of the hypotheses to data,
  - on a token-by-token basis
- Instead of comparing across experimental conditions, we compare each token directly against our hypotheses.
- **Benefits?** No assumption that all tokens in an experimental condition are uniform; not reliant on null hypothesis test
This talk: the Blackfoot LHL pitch contour

- Blackfoot (Algonquian; Frantz 2017)
- Stress can be on any syllable of the phonological phrase
- Pitch peak on stressed syllable (Miyashita & Weber 2020, Van Der Mark 2003, Weber 2020, 2016)
Interpretation and research question

- Assumptions: the pitch contour is due to a sequence of LHL targets, where
  - the H docks to a stressed syllable (here = 1st syllable)
  - final L% is a static boundary tone (Miyashita & Weber 2020)
  - the initial %L is introduced by the phonological phrase (Weber & Shaw 2022)
- Q: what phonological grammar determines location of the initial L?
Hypotheses: mobile vs. static boundary L tone

▶ Two competing hypotheses:
  1. static boundary tone hypothesis
  2. mobile boundary tone hypothesis

▶ different predictions for location of L in words with non-initial stress
▶ (same predictions for location of L in words with initial stress)
Contributions of the talk

**Methodological:** evaluate competing phonological hypotheses
- about *intonational grammars*
- which predict different *locations* of tonal targets
- rather than the presence/absence of a tonal target (as in Kawahara, Shaw & Ishihara 2021)

**Theoretical:** evaluate competing theoretical proposals of Blackfoot (Algonquian) intonational phonology on a token-by-token basis w.r.t. two hypotheses
- static boundary tone
- mobile boundary tone
Outline

Data collection

Analysis

Results

Discussion
Data collection
Participants

- \( n = 8 \) (4 male, 4 female)
- all are fluent Blackfoot speakers
  - Like most Blackfoot speakers (Genee & Junker 2018), the participants in our study use English in their daily lives, and can be characterized as English-dominant bilinguals.
  - Several participants are teachers of the Blackfoot language in a school setting.
- ages between 50 and 70 at time of recording
- residents of the Káínai Blackfoot reserve
Materials

- Part of a larger study with 52 nominal stems inflected for singular and plural.
- This study: one stem with 2nd syllable stress

<table>
<thead>
<tr>
<th>Word</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ma.ˈmɪn.n-i]</td>
<td>‘wing’</td>
</tr>
<tr>
<td>[ma.ˈmɪn.n-i²ts]</td>
<td>‘wings’</td>
</tr>
</tbody>
</table>

- /maˈmɪn:-/ ‘wing’ contains all sonorants = good for pitch tracking
- non-initial stress = phonological hypotheses predict different locations for initial %L
Initial instructions

- Speakers were asked to produce each word in a frame sentence twice.*
- Separate frame sentences for singular/plural.

(1) nitsííni’pa anní ___ matónni
    I.saw.it that(inan) ___ yesterday
    ‘I saw that ___ (inan., sg.) yesterday’

(2) nitsííni’pi anníístsi ___ matónni
    I.saw.them those(inan) ___ yesterday
    ‘I saw those ___ (inan., pl.) yesterday’

(*Some speakers preferred to create a new sentence for each word.*)
Picture prompts

- Pictures prompted either a singular or plural noun (= doubled image).
- Participants created a sentence on the fly.

**maminni ‘wing’**
Analysis
**Approach**

1. Pitch trajectories

2. Classifier

3. Classified tokens
1. Pitch trajectories

- f0 tracked with YAAPT (Zahorian & Hu 2008)
- we calculated each speaker’s average L and H pitch and average pitch at word onset, for use in constructing speaker-specific versions of our hypotheses.

3 (of 8) speakers excluded
- S01 and S06 had no tokens
- S04 had three tokens; one disfluent
- Included: S02, S03, S05, S07, S08
2. Classifier, Step 1: DCT

Represent $f_0$ trajectory as the sum of Cosines:

$$y(k) = w(k) \sum_{n=1}^{L} x(n) \cos \left( \frac{\pi (2n - 1)(k - 1)}{2L} \right)$$

$$k = 1, 2, \ldots, L$$

Where $L$ is the number of data samples and $x(n)$ is the trajectory to be modelled and:

$$w(k) = \begin{cases} 
\frac{1}{\sqrt{L}} & k = 1 \\
\sqrt{\frac{2}{L}} & 2 \leq k \leq L 
\end{cases}$$

2. Classifier, Step 1: Fit between real and simulated F0 using iDCT simulations using 6 DCT components are sufficient to explain greater than 0.95 of the variance for all speakers
2. Classifier, Step 2: iDCT

Simulate F0 trajectories from DCT components:

- **Static boundary tone**: \( y(k) \sim N(\mu(k), \sigma(k)) \)
- **Mobile boundary tone**: \( y(k) \sim N(\mu(k), \sigma(k)) \)

\[
x(n) = \sum_{n=1}^{L} w(k)y(k) \cos\left(\frac{\pi(2n - 1)(k - 1)}{2L}\right) \\
\]

\( n = 1, 2, \ldots L \)

Where \( L \) is the number of data samples and \( x(n) \) the trajectory to be simulated and:

\[
w(k) = \begin{cases} 
\frac{1}{\sqrt{L}} & k = 1 \\
\frac{2}{\sqrt{L}} & 2 \leq k \leq L 
\end{cases}
\]
2. Classifier, Step 3: Naive Bayes over DCT coefficients

\[ p(H|C_{o_1}, ..., C_{o_n}) = \frac{p(H) \times \prod_{i=1}^{n} p(C_{o_i}|H)}{\prod_{i=1}^{n} p(C_{o_i})} \]
3. Classified tokens: Hypothetical posterior probabilities

**H1:** Static boundary tone

**H2:** Mobile boundary tone
Results
Aggregated results

posterior probability of mobile boundary tone
SSANOVA of pitch trajectories

classified trajectories

competing hypotheses

posterior probability of mobile boundary tone
SSANOVA of pitch trajectories

classified trajectories

competing hypotheses

posterior probability of mobile boundary tone
Summary of results

- Results support the ‘mobile boundary tone’ hypothesis
  - Despite substantial phonetics variability
  - All tokens were classified as the ‘mobile boundary tone’ hypothesis
- Theoretical claim: boundary tones are tones that *co-occur* with a prosodic boundary
  - Potentially dock to a non-edge syllable
  - Not lexical tones, which gives the appearance of being ‘mobile’
Discussion
Converging evidence

▶ Converging evidence with results from a previous study using standard LabPhon methods (Weber & Shaw 2022)
▶ Compared pitch slope and L timing across words which vary by stress
▶ However, small data (N = 94) makes results statistically marginal.
Comparing the two methods

- Weber and Shaw (2022) ‘standard’ approach:
  - based on pitch slope and timing of L (sparse phonetic data)
  - not enough data to draw reliable conclusions with standard statistical methods

- Simulation and classification methods:
  - based on continuous pitch contour (rich phonetic data)
  - continuous posterior probabilities rather than null hypothesis tests
  - token-by-token evaluation reveals potential for variation within conditions

In this case, the two methods both converge on the same result.
Expanding the empirical domain of the methodology

- Previous studies using these methods:
  - presence/absence of a vowel (Shaw & Kawahara 2018)
  - presence/absence of pitch accent (Kawahara, Shaw & Ishihara 2021)

- This study:
  - two separate pitch contours (differentiated by location of L pitch target)
  - derived from competing phonological hypotheses (Weber & Shaw 2022)
Data requirements of the methodology

- Less data-intensive than the standard approach
  - number of tokens can be very small!
  - caveat: need to estimate variability
- Potentially enables robust evaluation of theoretical hypotheses from smaller and less controlled datasets, e.g.
  - languages with few speakers
  - data sources like narratives, conversations
  - etc.
Acknowledgements

► Thanks to Francis First Charger, John Heavy Shield, Kim Black Water, Leo Fox, Natalie Creighton, Peter Weasel Mocassin, and others who have shared their language with us and literally made this project possible. Nitsíkohtaahsi’taki!

► Phonology Reading Group at Yale (esp. Sarah Babinsky, Sammy Anderson, Chelsea Sanker, Roslyn Burns for helpful comments!)
References


