Comparing cues

A mixed methods study of intonation unit boundaries in three typologically diverse languages

Naomi Peck\textsuperscript{1}, Kirsten Culhane\textsuperscript{1}, Maria Vollmer\textsuperscript{1, 2}

\textsuperscript{1}University of Freiburg \textsuperscript{2}ARC Centre of Excellence for the Dynamics of Language
We gratefully acknowledge the Kera'a, Waima'a and Warlpiri people who provided the material used in this presentation. We acknowledge that the Warlpiri data was recorded on Warlpiri and Gurindji land and we pay our respects to their elders past, present and emerging.
Background

In this presentation, we investigate:

- which boundary cues are most frequent
- which boundary cues are most salient
- what motivates cue choice

in three understudied languages.
Background

Motivation for the study

- Claims for universality of intonation unit boundary cues, like pauses, pitch resets, final lengthening and initial rushes (Himmelmann et al. 2018)
- Variation of the relative importance of intonational boundary cues on a language-specific basis (Izre’el and Mettouchi 2015: 23)
- Different realisations of most common cues across languages (Himmelmann et al. 2018: 239)
- Less systematic boundary cues such as creaky voice reinforcing the perception of prosodic boundaries (Wagner and Watson 2010)
Background

The languages we investigate have different typological profiles:

<table>
<thead>
<tr>
<th>Language</th>
<th>Family</th>
<th>Region</th>
<th>Lexical Tone</th>
<th>Lexical Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waima'a</td>
<td>Austronesian</td>
<td>Timor-Leste</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Warlpiri</td>
<td>Pama–Nyungan</td>
<td>Northern Australia</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>Kera'a</td>
<td>Tibeto-Burman</td>
<td>NE India</td>
<td>✓</td>
<td>?</td>
</tr>
</tbody>
</table>
Hypothesis

We hypothesise:

1. some cues are more salient than others
2. speakers of each selected language will differ in what cues they use most frequently
3. that the choice of cues will be affected by the discourse organisation of an utterance
4. that cue choice will be affected by the typological profile of each language
Methods

Using naturalistic field data, we segmented texts into intonation units, followed by a manual annotation in Praat for presence of boundary cues.

<table>
<thead>
<tr>
<th>Language</th>
<th>Minutes</th>
<th>Speakers</th>
<th>Genres</th>
<th>Text no.</th>
<th>IU no.</th>
<th>Sources</th>
<th>Coder</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kera’a</td>
<td>15</td>
<td>2 (2M)</td>
<td>Narrative</td>
<td>3</td>
<td>609</td>
<td>Own fieldwork (2020, Naomi)</td>
<td>Naomi, Sarah Stolle</td>
</tr>
</tbody>
</table>
Cues

We coded for cues of IU boundaries cited in the literature:

- Pitch resets (Himmelmann et al. 2018)
- Pauses (Himmelmann et al. 2018)
- Final lengthening (of vowels) (Cruttenden 1997, Fletcher 2010)
- Initial rush (anacrusis) (Cruttenden 1997, Fletcher 2010)
- Creaky voice phenomena (cf. Davidson 2020)

We also annotated for other cues we found in these languages which aren’t as commonly cited:

- Tonal parallelism (Croft 2007)
- Final rushes (Fletcher 2010)
- Final devoicing (Kilbourn-Ceron and Sonderegger 2018)
- Initial lengthening (of consonants) (Keating et al. 2003)
Rarer cues: Tonal parallelism

“We keep chickens, pigs, mithuns…”
Rarer cues: Final rushes

‘The goatherd…’
Rarer cues: Final devoicing

‘...they became 7 women’
Rarer cues: Initial lengthening of consonants

‘When returning from there...’
Results

We hypothesise:

1. some cues are more salient than others
2. speakers of each selected language will differ in what cues they use most frequently
3. that the choice of cues will be affected by the discourse organisation of an utterance,
4. that cue choice will be affected by the typological profile of each language
## Intercoder Reliability

<table>
<thead>
<tr>
<th>Measure</th>
<th>Gwet’s AC1</th>
<th>Brennan-Prediger</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitch Reset</td>
<td>Very Good</td>
<td>Good</td>
</tr>
<tr>
<td>Pause</td>
<td>Very Good</td>
<td>Very Good</td>
</tr>
<tr>
<td>Initial Lengthening</td>
<td>Very Good</td>
<td>Very Good</td>
</tr>
<tr>
<td>Final Lengthening</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>Initial Rush</td>
<td>Good</td>
<td>Good</td>
</tr>
<tr>
<td>Final Rush</td>
<td>Very Good</td>
<td>Very Good</td>
</tr>
<tr>
<td>Tonal Parallelism</td>
<td>Very Good</td>
<td>Good</td>
</tr>
<tr>
<td>Creaky Voice</td>
<td>Good</td>
<td>Moderate</td>
</tr>
<tr>
<td>Final Devoicing</td>
<td>Very Good</td>
<td>Good</td>
</tr>
</tbody>
</table>
Differences in Cue Realisation

We found that some cues were encoded differently by speakers of different languages.

For example, creaky voice:
Relative importance of cues

Pitch resets and pauses were the most frequent boundary cues in our study. The relatively high reliability of coding also supports the fact that these features are cross-linguistically the most salient cues.

However, the relative importance of cues differs from language to language.

<table>
<thead>
<tr>
<th></th>
<th>Warlpiri</th>
<th>Waima’a</th>
<th>Kera’a</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pauses</td>
<td>Pitch Reset</td>
<td>Pitch Reset</td>
</tr>
<tr>
<td>2</td>
<td>Pitch Reset</td>
<td>Pauses</td>
<td>Pauses</td>
</tr>
<tr>
<td>3</td>
<td>Creaky Voice</td>
<td>Final Lengthening</td>
<td>Final Lengthening</td>
</tr>
<tr>
<td>4</td>
<td>Final Devoicing</td>
<td>Final Rush</td>
<td>Creaky Voice</td>
</tr>
<tr>
<td>Cue</td>
<td>Warlpiri</td>
<td>Cue</td>
<td>Waima’a</td>
</tr>
<tr>
<td>-------------</td>
<td>----------</td>
<td>--------------</td>
<td>---------</td>
</tr>
<tr>
<td>Pauses</td>
<td>93%</td>
<td>Pitch Reset</td>
<td>94%</td>
</tr>
<tr>
<td>Pitch Reset</td>
<td>87%</td>
<td>Pauses</td>
<td>84%</td>
</tr>
<tr>
<td>Creaky</td>
<td>25%</td>
<td>Fin. Length.</td>
<td>20%</td>
</tr>
<tr>
<td>Devoicing</td>
<td>25%</td>
<td>Fin. Rush</td>
<td>9%</td>
</tr>
<tr>
<td>Fin. Length.</td>
<td>18%</td>
<td>Parallelism</td>
<td>9%</td>
</tr>
<tr>
<td>Fin. Rush</td>
<td>13%</td>
<td>Devoicing</td>
<td>7%</td>
</tr>
<tr>
<td>Parallelism</td>
<td>13%</td>
<td>Init. Rush</td>
<td>5%</td>
</tr>
<tr>
<td>Init. Rush</td>
<td>9%</td>
<td>Creaky</td>
<td>5%</td>
</tr>
<tr>
<td>Init. Length.</td>
<td>2%</td>
<td>Init. Length.</td>
<td>4%</td>
</tr>
</tbody>
</table>
Results

We hypothesise:

1. some cues are more salient than others
2. speakers of each selected language will differ in what cues they use most frequently
3. that the choice of cues will be affected by the discourse organisation of an utterance
4. that cue choice will be affected by the typological profile of each language
Discourse Organisation

● Initial and mostly impressionistic tendencies

● Discourse organisation effects on:
  ○ Tonal parallelism
  ○ Final lengthening and devoicing

● Tendencies either universal or language-specific
Lists and Tonal Parallelism

‘Actually, (the fire wattle; topic of the story) are wurruru for Japanangka, Japangardi and Napanangka.’ (Morton 2009b) (wurruru = egocentric relational term for other half of patrimoiety)
Paratones, Final Lengthening and Devoicing

‘And so, they… the wind and the clouds, everything has a place in that tree.’

Only time final devoicing occurs
No final lengthening
Results

We hypothesise:

1. some cues are more salient than others
2. speakers of each selected language will differ in what cues they use most frequently
3. that the choice of cues will be affected by the discourse organisation of an utterance
4. that cue choice will be affected by the typological profile of each language
Interpreting Cue Choice

The preferences for certain boundary cues can be linked to the typological profile of each language.

- **Lexical tone in Kera’a**
  - Less final devoicing
  - More creaky voice

- **CV(C)V phonological word template in Waima’a**
  - Final devoicing only occurs in CV final words

- ‘Intonation only’ system in Waima’a
  - Pitch reset as the main melodic cue

- **V-final phonological words in Warlpiri**
  - Both final devoicing and final lengthening as frequent cues (but could be epiphenomenal result of language change)
Revisiting the Rarer Cues

We hypothesise:

1. some cues are more salient than others
2. speakers of each selected language will differ in what cues they use most frequently
3. that the choice of cues will be affected by the discourse organisation of an utterance
4. that cue choice will be affected by the typological profile of each language
Creaky Voice

Literature has previously downplayed the role of creaky voice as a melodic boundary cue for IUs (e.g. Wagner and Watson 2010:910, Himmelmann et al. 2018:214).

However, our results suggest it is just as frequent (if not more frequent) of a strategy to mark IU boundaries as initial rushes in our data, a boundary cue which has traditionally received more attention.
<table>
<thead>
<tr>
<th>Cue</th>
<th>Warlpiri</th>
<th>Cue</th>
<th>Waima’a</th>
<th>Cue</th>
<th>Kera’a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pauses</td>
<td>93%</td>
<td>Pitch Reset</td>
<td>94%</td>
<td>Pitch Reset</td>
<td>91%</td>
</tr>
<tr>
<td>Pitch Reset</td>
<td>87%</td>
<td>Pauses</td>
<td>84%</td>
<td>Pauses</td>
<td>82%</td>
</tr>
<tr>
<td>Creaky</td>
<td>25%</td>
<td>Fin. Length.</td>
<td>20%</td>
<td>Fin. Length.</td>
<td>45%</td>
</tr>
<tr>
<td>Devoicing</td>
<td>25%</td>
<td>Fin. Rush</td>
<td>9%</td>
<td>Creaky</td>
<td>44%</td>
</tr>
<tr>
<td>Fin. Length.</td>
<td>18%</td>
<td>Parallelism</td>
<td>9%</td>
<td>Init. Rush</td>
<td>20%</td>
</tr>
<tr>
<td>Fin. Rush</td>
<td>13%</td>
<td>Devoicing</td>
<td>7%</td>
<td>Parallelism</td>
<td>11%</td>
</tr>
<tr>
<td>Parallelism</td>
<td>13%</td>
<td>Init. Rush</td>
<td>5%</td>
<td>Fin. Rush</td>
<td>5%</td>
</tr>
<tr>
<td>Init. Rush</td>
<td>9%</td>
<td>Creaky</td>
<td>5%</td>
<td>Devoicing</td>
<td>2%</td>
</tr>
<tr>
<td>Init. Length.</td>
<td>2%</td>
<td>Init. Length.</td>
<td>4%</td>
<td>Init. Length.</td>
<td>0%</td>
</tr>
</tbody>
</table>
Creaky Voice

The relatively high frequency of creaky voice could be linked to a few different factors:

- Creaky voice associated with lexical tone realisation in Kera’a
- Older speakers speaking with creak in Warlpiri data
- Boundary strength (e.g. Henton and Bladon 1988, Kuang 2018)

But this can’t ‘explain it all away’, and certainly not the variation between languages either.
Expanding the typology

Are there other relevant rhythmic boundary cues apart from initial rushes and final lengthening?
Expanding the typology

- Final rushes used 5~10% of the time
  - More common than initial rushes in Warlpiri and Waima’a
Expanding the typology

- Final rushes used 5~10% of the time
  - More common than initial rushes in Warlpiri and Waima’a

- Initial lengthening is rare (0~4%)
  - cf. White et al. 2020:12 for possible explanations
Expanding the typology

- Final rushes used 5~10% of the time
  - More common than initial rushes in Warlpiri and Waima’a

- Initial lengthening is rare (0~4%)
  - cf. White et al. 2020:12 for possible explanations

- Final rushes and initial lengthening more reliably coded than initial rushes and final lengthening
Expanding the typology

We can capture the distribution of these rhythmic cues like so:

<table>
<thead>
<tr>
<th></th>
<th>Initial</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rush</td>
<td>Initial rush (anacrusis)</td>
<td>Final rush</td>
</tr>
<tr>
<td>Lengthening</td>
<td>Initial lengthening</td>
<td>Final lengthening</td>
</tr>
</tbody>
</table>
Limitations and Future Research

We hypothesise:

1. some cues are more salient than others
2. speakers of each selected language will differ in what cues they use most frequently
3. that the choice of cues will be affected by the discourse organisation of an utterance
4. that cue choice will be affected by the typological profile of each language
Limitation 1: Dataset

- Small datasets (438-609 IUs / 2-5 speakers per language)
Limitation 1: Dataset

- Small datasets (438-609 IUs / 2-5 speakers per language)
- Rough control for genre
Limitation 2: Coding

- Perceptual coding by non-native speakers
  - But with help of Praat
  - But with help of the intercoder reliability test
  - But fairly consistent coding by non-native speakers according to Himmelmann et al. (2018)
Limitation 2: Coding

- Perceptual coding with non-native speakers
- Circular coding: Segmentation $\leftrightarrow$ Boundary Cues

Preconceived notions of which boundary cues are important may have influenced our segmentation of IUs, which would then influence which cues we code.

- Only salient cues
- Adjusted boundary cues for second round of coding
Limitation 3: Underdescribed languages

We are dealing with incomplete phonological descriptions and language change.
What we perceive as boundary cues may be an epiphenomenon of something else.

→ Consequence of looking at understudied and changing languages
→ That’s why it’s worth it!
Future Research

- Phonological profiles and language change
Future Research

- Phonological profiles and language change
- Expanded dataset (more IUs, speakers, genres, languages)
Future Research

- Phonological profiles and language change
- Expanded dataset (more IUs, speakers, genres, languages)
- Better control for speaker / genre effect
Future Research

- Phonological profiles and language change
- Expanded dataset (more IUs, speakers, genres, languages)
- Better control for speaker / genre effect
- More fine-grained quantitative and qualitative analyses
  - Reference tracking / coding of paratones / coding of topic and focus
  - Co-occurrence of cues
Summary

- Even when languages use the same cues, they use them differently
  - Cue preference
  - Cue realisation
Summary

- Even when languages use the same cues, they use them differently
  - Cue preference
  - Cue realisation

- Discourse organisation and typological profiles affect cue choice
Summary

- Even when languages use the same cues, they use them differently
  - Cue preference
  - Cue realisation

- Discourse organisation and typological profiles affect cue choice

- Lesser-studied cues are important
  - Creaky voice
  - Final rushes
Summary

- Even when languages use the same cues, they use them differently
  - Cue preference
  - Cue realisation

- Discourse organisation and typological profiles affect cue choice

- Lesser-studied cues are important
  - Creaky voice
  - Final rushes

- The benefits outweigh the costs of working with lesser-studied languages
Thank you for listening!
We would like to thank

Kera’a, Waima’a and Warlpiri speaking communities

Jane Simpson (ANU), David Nash (ANU), and Carmel O’Shannessy (ANU)

Sarah Stolle

Funding from Deutsche Forschungsgemeinschaft (DFG)
Project number 406074683

Support from ARC Centre of Excellence for the Dynamics of Language for Warlpiri
Project ID: CE140100041
References

References


We would like to thank

Kera’a, Waima’a and Warlpiri speaking communities

Jane Simpson (ANU), David Nash (ANU), and Carmel O’Shan Nessy (ANU)

Sarah Stolle

Funding from Deutsche Forschungsgemeinschaft (DFG)
Project number 406074683

Support from ARC Centre of Excellence for the Dynamics of Language for Warlpiri
Project ID: CE140100041
We would like to thank

Kera’a, Waima’a and Warlpiri speaking communities

Jane Simpson (ANU), David Nash (ANU), and Carmel O’Shannessy (ANU)

Sarah Stolle

Funding from Deutsche Forschungsgemeinschaft (DFG)
Project number 406074683

Support from ARC Centre of Excellence for the Dynamics of Language for Warlpiri
Project ID: CE140100041
Methods

- Segmentation into intonation units, followed by manual annotation in Praat
- 4 coders
  - 2 native speakers of Australian English (Kirsten + Naomi)
  - 2 native speakers of German (Maria + Sarah)

<table>
<thead>
<tr>
<th>Language</th>
<th>Coding Details</th>
<th>Kirsten</th>
<th>Sarah</th>
<th>Everyone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waima’a</td>
<td>9 minutes</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>coded by Kirsten</td>
<td></td>
<td>5 minutes coded by Sarah</td>
<td>+</td>
</tr>
<tr>
<td>Warlpiri</td>
<td>13 minutes</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>coded by Maria</td>
<td></td>
<td>5 minutes coded by Sarah</td>
<td>+</td>
</tr>
<tr>
<td>Kera’a</td>
<td>9 minutes</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>coded by Naomi</td>
<td></td>
<td>5 minutes coded by Sarah</td>
<td>+</td>
</tr>
</tbody>
</table>
Discourse organisation and Rushes

- In Kera’a, rushes tend to occur with
  a. Discourse regulation (hesitation, repairs, um/so/yeah)
  b. Function words
  c. Given or accessible ideas (after Chafe)
<table>
<thead>
<tr>
<th>Cue</th>
<th>Warlpiri</th>
<th>Cue</th>
<th>Waima’a</th>
<th>Cue</th>
<th>Kera’a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pauses</td>
<td>439/473</td>
<td>Pitch Reset</td>
<td>405/433</td>
<td>Pitch Reset</td>
<td>550/606</td>
</tr>
<tr>
<td>Pitch Reset</td>
<td>411/471</td>
<td>Pauses</td>
<td>368/438</td>
<td>Pauses</td>
<td>499/609</td>
</tr>
<tr>
<td>Creaky</td>
<td>123/485</td>
<td>Fin. Length.</td>
<td>86/438</td>
<td>Fin. Length.</td>
<td>272/609</td>
</tr>
<tr>
<td>Devoicing</td>
<td>122/485</td>
<td>Fin. Rush</td>
<td>39/438</td>
<td>Creaky</td>
<td>266/609</td>
</tr>
<tr>
<td>Fin. Length.</td>
<td>85/485</td>
<td>Parallelism</td>
<td>39/438</td>
<td>Init. Rush</td>
<td>119/609</td>
</tr>
<tr>
<td>Fin. Rush</td>
<td>61/485</td>
<td>Devoicing</td>
<td>30/438</td>
<td>Parallelism</td>
<td>67/609</td>
</tr>
<tr>
<td>Parallelism</td>
<td>64/485</td>
<td>Init. Rush</td>
<td>24/438</td>
<td>Fin. Rush</td>
<td>32/609</td>
</tr>
<tr>
<td>Init. Rush</td>
<td>45/485</td>
<td>Creaky</td>
<td>21/438</td>
<td>Devoicing</td>
<td>9/609</td>
</tr>
<tr>
<td>Init. Length.</td>
<td>9/485</td>
<td>Init. Length.</td>
<td>18/438</td>
<td>Init. Length.</td>
<td>0/609</td>
</tr>
</tbody>
</table>
# Intercoder Reliability

<table>
<thead>
<tr>
<th>Measure</th>
<th>AC1</th>
<th>Brennan-Prediger</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitch Reset</td>
<td>0.821</td>
<td>0.704</td>
</tr>
<tr>
<td>Pause</td>
<td>0.904</td>
<td>0.872</td>
</tr>
<tr>
<td>Initial Lengthening</td>
<td>0.944</td>
<td>0.896</td>
</tr>
<tr>
<td>Final Lengthening</td>
<td>0.517</td>
<td>0.481</td>
</tr>
<tr>
<td>Initial Rush</td>
<td>0.767</td>
<td>0.649</td>
</tr>
<tr>
<td>Final Rush</td>
<td>0.890</td>
<td>0.809</td>
</tr>
<tr>
<td>Tonal Parallelism</td>
<td>0.818</td>
<td>0.716</td>
</tr>
<tr>
<td>Creaky Voice</td>
<td>0.641</td>
<td>0.484</td>
</tr>
<tr>
<td>Final Devoicing</td>
<td>0.859</td>
<td>0.771</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Range</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8 - 1.0</td>
<td>Very Good</td>
</tr>
<tr>
<td>0.6 - 0.8</td>
<td>Good</td>
</tr>
<tr>
<td>0.4 - 0.6</td>
<td>Moderate</td>
</tr>
<tr>
<td>0.2 - 0.4</td>
<td>Fair</td>
</tr>
<tr>
<td>-1.0 - 0.2</td>
<td>Poor</td>
</tr>
</tbody>
</table>

(Altman 1991)