Methods for detecting prosodic structure

Microparametric variation in prosodic structure: case studies from Algonquian

Natalie Weber*, Antti Arppe†, Ksenia Bogomolets‡, Andrew Cowell§, Rose-Marie Déchaineǁ, Christopher Hammerlyǁ, Sarah Murrayǁ, Katherine Schmirler†, Rachel Vogel*

*Yale University, †University of Alberta, ‡University of Auckland, §University of Colorado, Boulder, ǁUniversity of British Columbia, ǁCornell University
Goals of the project

● Prosodic structure is related to syntactic structure, but can mismatch (cf. Kaisse 1985; Nespor and Vogel 2007 [1986]; Selkirk 1986)

● Variation in prosodic structure may derive from
  ○ different syntactic structures
  ○ different phonological grammars

● Solution? Languages with similar morphosyntactic template!
  ○ examine parametric variation in prosodic structure
  ○ very likely to be specific to prosody, not syntax*

● **Micro**parametric = within a single language family.

* future work: how does syntax map to the morphosyntactic template?
Algonquian family

[Ish ishwar, CC BY 2.0 <https://creativecommons.org/licenses/by/2.0/deed.en>, via Wikimedia Commons]
Current working group members

Plains Algonquian

- Blackfoot  Natalie Weber (Yale)
- Cheyenne  Sarah Murray (Cornell); Rachel Vogel (Yale)
- Arapaho  Andrew Cowell (Boulder); Ksenia Bogomolets (Auckland)

Central Algonquian

- Plains Cree  Rose-Marie Déchaine (UBC); Antti Arppe (Alberta)
  Katherine Schmirler (Alberta/Lethbridge)
- Ojibwe  Chris Hammerly (UBC)

Eastern Algonquian

- Passamaquoddy  (in progress)
Why Algonquian?

- **Polysynthetic** (Baker 1996; Duponceau 1819; Mattissen 2004; Nichols 1986, 1992; a.o.)
  - strongly head-marking
  - extensive agglutinative morphology
  - multiple “lexical” morphemes or roots within a morphological “word”, incl. adjuncts (Mathieu, Fry, & Barrie 2017; Newell & Piggott 2014; Piggott & Newell 2006; Piggott & Travis 2013)

- Theories of prosody-syntax correspondence remain poorly tested on polysynthetic languages (Elfner 2018; but see Miller 2018; Miller & Sande 2021; Weber 2020, 2021, 2022; and case-studies in Bogomolets & van der Hulst *to appear*)
Why now?

- Long history of diachronic linguistic analysis (cf. Aubin 1975; Bloomfield 1925, 1946; Goddard 1979; Hewson 1993; Hockett 1942; Michelson 1935; Miller 1959; Pentland 1979; Siebert 1941, 1975; Silver 1960; Voegelin 1941)
- Less on synchronic phonological generalizations
- Recent resources:
  - descriptive grammars and dictionaries for all languages in the study
  - large textual and spoken corpora for many languages!
  - multiple data structures to probe for prosodic structure
Roadmap

Four 15-minute talks (including discussant)

1. Methods for detecting prosodic structure *(current)*
2. Phonological alternations at morphological edges
3. Determining prosodic constituency from morpho-phonological generalizations
4. Discussant: Emily Elfner (York University)

Q&A session: 30 minutes
Overview
Prosodic structure

- “Interface” categories: derive from, but are not isomorphic to, syntax.

PPh = Phonological Phrase
CG = Composite Group
PWd = Prosodic Word
PStem = Prosodic Stem

“prosodic” in this talk refers to interface categories, not metrical structure or stress

Indirect Reference Hypothesis

- Prosodic units (not syntactic units) define the domains for phonological processes and generalizations.
- e.g. primary word stress and phrase level stress readjustment in English

This is [the cat that caught [the rat that stole [the cheese]]]

(Chomsky & Halle 1968: 372; broad IPA added)
Algonquian template [Plains Cree]

- aimed at language comparison and reconstruction of the Algonquian family
- template slots named by their positions
- “preverbs” = mixed bag, including event modifiers (adjuncts)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pim</td>
<td>–ohtê</td>
<td>–w</td>
</tr>
<tr>
<td>sâpo–</td>
<td>pim</td>
<td>–ohtê</td>
<td>–w</td>
</tr>
<tr>
<td>ki–</td>
<td>sâpo–</td>
<td>pim</td>
<td>–ohtâ</td>
</tr>
</tbody>
</table>

(For Plains Cree see Wolvengrey 2001; for Algonquian see Bloomfield 1946: 111, #269; Goddard 1990)
Mapping the Algonquian template to phrasal syntax

\[ \sqrt{\text{pim}} \quad \text{–ohtê} \quad \text{–w} \]
\[ \sqrt{\text{along}} \quad \text{–walk.AI} \quad \text{–3} \]
‘s/he walks along’

\[ \sqrt{\text{sâpo}} \quad \text{–hâ} \quad \text{–w} \]
\[ \sqrt{\text{through}} \quad \text{–fly.AI} \quad \text{–3} \]
‘s/he flies through’

\[ \sqrt{\text{sâpo–}} \quad \sqrt{\text{pim}} \quad \text{–ohtê} \quad \text{–w} \]
\[ \sqrt{\text{through–}} \quad \sqrt{\text{along}} \quad \text{–walk.AI} \quad \text{–3} \]
‘s/he walks past’

(Bruening 2001:122; Brittain 2003; Hirose 2003; Branigan et al. 2005; Piggott and Newell 2006; Mathieu 2007; Slavin 2012)
Central questions for this project

1. How does morphological template correspond to prosodic structure?
2. How does prosodic structure vary across languages?

H1 \(((\text{prefixes}–\text{preverb}–\text{initial}–\text{final}–\text{suffixes})_{\text{PStem}})_{\text{PWD}}\)

H2 \((\text{prefixes}–\text{preverb}–(\text{initial}–\text{final}–\text{suffixes})_{\text{PStem}})_{\text{PWD}}\)

H3 \((\text{prefixes}–(\text{preverb})_{\text{PStem}}(\text{initial}–\text{final}–\text{suffixes})_{\text{PStem}})_{\text{PWD}}\)
Comparison to previous work

● Previous work on Algonquian prosodic structure focuses on the mapping of syntax to prosodic structure (Branigan et al. 2005; Mathieu, Fry, & Barrie 2017; Newell & Piggott 2014; Piggott & Newell 2006; Piggott & Travis 2013; Russell 1999; Russell & Reinholtz 1997)

● In contrast, we focus on building up the prosodic structure from basic phonological generalizations
  ○ generalizations (distribution, phonotactics, alternations)
  ○ hold across prosodic constituents or at prosodic boundaries
  ○ templatic positions within the same prosodic unit should pattern alike
Methodological toolkit
Methodological toolkit

Data comes from:

- **Annotated textual corpora**
  - Systematic, based on dictionary entries
  - But relies on orthography

- **Speech corpora**
  - Unsystematic, but natural speech
  - Lends itself to computational methods
Annotated text-based resources

- Types of documentation differ across languages (wordlists, grammars, dictionaries)
- Dictionaries often stem-based, with stem-internal morphology unmarked
  - creating annotations within spreadsheets
  - able to build paradigms on the fly
  - sort by any position within template
- (Some confirmation of previous observations; some new observations.)

It is a nice house, is a nice room; it (a dwelling) is in good order.

**minwate** 0s ind; **minwateg** 0s conj; **menwateg** 0s ch-conj; Stem: */minwate-/

### Word Parts

- **minwate** /minwate-/: /minw-/ good; /-ate/ it is or is in an interior space, a room, a house
<table>
<thead>
<tr>
<th>Initial UR</th>
<th>Initial Gloss</th>
<th>Final UR</th>
<th>Final Gloss</th>
<th>Morphemic analysis (surface)</th>
<th>Original Word</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>madwe-</td>
<td>heard, audible</td>
<td>-aasin</td>
<td>it is blown by the madwe-bag-aasin</td>
<td>madwebagaasin</td>
<td>the leaves can be heard blowing in the</td>
<td>142</td>
</tr>
<tr>
<td>madwe-</td>
<td>heard, audible</td>
<td>-am</td>
<td>act on h/ by mou madwe_-m</td>
<td>madwem</td>
<td>chew it (animate) audibly</td>
<td>143</td>
</tr>
<tr>
<td>madwe-</td>
<td>heard, audible</td>
<td>-and</td>
<td>act on it by mout madwe_-nd-an</td>
<td>madwendan</td>
<td>chew it audibly</td>
<td>144</td>
</tr>
<tr>
<td>mane-</td>
<td>scarce, lack, need</td>
<td>-izi</td>
<td>s/he, it (animate) mane-zi-n</td>
<td>manezin</td>
<td>s/he is in need of, is short of (it)</td>
<td>145</td>
</tr>
<tr>
<td>mang-</td>
<td>big</td>
<td>-izi</td>
<td>s/he, it (animate) mang-ade-zi</td>
<td>mangadezi</td>
<td>s/he is wide</td>
<td>146</td>
</tr>
<tr>
<td>minw-</td>
<td>good</td>
<td>-aabi</td>
<td>s/he looks, has s'minw-aabi</td>
<td>minwaabi</td>
<td>s/he has good eyesight, sees well</td>
<td>147</td>
</tr>
<tr>
<td>minw-</td>
<td>good</td>
<td>-ate</td>
<td>it is or is in an in minw-ate</td>
<td>minwate</td>
<td>it is a nice house, is a nice room; it (a</td>
<td>148</td>
</tr>
<tr>
<td>minw-</td>
<td>good</td>
<td>-bii</td>
<td>s/he or it (animate) min-o-bii</td>
<td>minobii</td>
<td>s/he drinks and is merry</td>
<td>149</td>
</tr>
<tr>
<td>minw-</td>
<td>good</td>
<td>-enim</td>
<td>act by thought or minw-enim</td>
<td>minwenim</td>
<td>like h/</td>
<td>150</td>
</tr>
<tr>
<td>minw-</td>
<td>good</td>
<td>-gaabawi</td>
<td>h/ stands</td>
<td>mino-gaabawi</td>
<td>s/he stands well; s/he is in good stand</td>
<td>151</td>
</tr>
<tr>
<td>minw-</td>
<td>good</td>
<td>-lalaw</td>
<td>hear h/</td>
<td>mino-taw</td>
<td>like hearing h/, like how s/he sounds</td>
<td>152</td>
</tr>
<tr>
<td>misaw-</td>
<td>desire</td>
<td>-enim</td>
<td>act by thought or misaw-enim</td>
<td>misawenim</td>
<td>want, desire h/</td>
<td>153</td>
</tr>
<tr>
<td>misaw-</td>
<td>desire</td>
<td>-n</td>
<td>see it</td>
<td>misaw-i-n-an</td>
<td>wish to have it</td>
<td>154</td>
</tr>
<tr>
<td>misaw-</td>
<td>desire</td>
<td>-naw</td>
<td>see h/</td>
<td>misaw-i-naw</td>
<td>wish to have h/; be envious of h/</td>
<td>155</td>
</tr>
<tr>
<td>miskw-</td>
<td>red</td>
<td>-aa</td>
<td>it is in a state or misk-o-bag-aa</td>
<td>miskobagaa</td>
<td>there are red leaves</td>
<td>156</td>
</tr>
<tr>
<td>#</td>
<td>Initial UR</td>
<td>Initial Gloss</td>
<td>Final UR</td>
<td>Final Gloss</td>
<td>Morphemic analysis (surface)</td>
<td>Original Word</td>
</tr>
<tr>
<td>---</td>
<td>------------</td>
<td>---------------</td>
<td>--------------</td>
<td>-------------</td>
<td>-----------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>89</td>
<td>agaas-</td>
<td>small</td>
<td>-ate</td>
<td>it is or is in an int agaas-ate</td>
<td>agaasate</td>
<td>it is a small house or room</td>
</tr>
<tr>
<td>90</td>
<td>biin-</td>
<td>clean</td>
<td>-ate</td>
<td>it is or is in an int biin-ate</td>
<td>biinate</td>
<td>it is a clean house or room</td>
</tr>
<tr>
<td>91</td>
<td>dak-</td>
<td>cool, chilly, cold</td>
<td>-ate</td>
<td>it is or is in an int dak-ate</td>
<td>dakate</td>
<td>it (a room or house) is cold</td>
</tr>
<tr>
<td>92</td>
<td>giizhoo-</td>
<td>warm</td>
<td>-ate</td>
<td>it is or is in an int giizhoo- _te</td>
<td>giizhoote</td>
<td>it is warm inside</td>
</tr>
<tr>
<td>93</td>
<td>minw-</td>
<td>good</td>
<td>-ate</td>
<td>it is or is in an int minw-ate</td>
<td>minwate</td>
<td>it is a nice house, is a nice room; it (a</td>
</tr>
<tr>
<td>94</td>
<td>waase-</td>
<td>clear, light</td>
<td>-ate</td>
<td>it is or is in an int waase- _te</td>
<td>waasete</td>
<td>it is or gets bright inside</td>
</tr>
<tr>
<td>95</td>
<td>dwaa-</td>
<td>going through the</td>
<td>-bide</td>
<td>it moves without dwaa-bid-e</td>
<td>dwaabide</td>
<td>it drives through ice</td>
</tr>
<tr>
<td>96</td>
<td>akw-</td>
<td>a certain length, sx</td>
<td>-bii</td>
<td>s/he or it (animal ak-o-bii)</td>
<td>akobii</td>
<td>s/he is so far into the water</td>
</tr>
<tr>
<td>97</td>
<td>akw-</td>
<td>a certain length, sx</td>
<td>-bii</td>
<td>liquid, water, slur ak-o-bii</td>
<td>akobii</td>
<td>it has water or liquid in it so high</td>
</tr>
<tr>
<td>98</td>
<td>deb-</td>
<td>enough, adequate</td>
<td>-bii</td>
<td>it is or does relat deb-i-bii</td>
<td>debibii</td>
<td>it fits (of liquid); it adequately holds (sc</td>
</tr>
<tr>
<td>99</td>
<td>gaw-</td>
<td>prostrate, down ar</td>
<td>-bii</td>
<td>s/he or it (animal gaw-i-bii)</td>
<td>gawibii</td>
<td>s/he falls down drunk</td>
</tr>
<tr>
<td>100</td>
<td>minw-</td>
<td>good</td>
<td>-bii</td>
<td>s/he or it (animal min-o-bii)</td>
<td>minobii</td>
<td>s/he drinks and is merry</td>
</tr>
<tr>
<td>101</td>
<td>naad-</td>
<td>fetch, go get, appr</td>
<td>-bii</td>
<td>s/he or it (animal naad-oob-ii)</td>
<td>naadoobii</td>
<td>s/he goes to get water or other liquid;</td>
</tr>
<tr>
<td>102</td>
<td>naad-</td>
<td>fetch, go get, appr</td>
<td>-bii</td>
<td>s/he or it (animal naaz-i-bii)</td>
<td>naazibii</td>
<td>s/he goes down to the water, goes aft</td>
</tr>
<tr>
<td>103</td>
<td>dwaa-[NW]</td>
<td>going through the</td>
<td>-bii [NW]</td>
<td>liquid, water, slur dwaa-́-i-bii</td>
<td>dwaa'ibii</td>
<td>s/he makes a water hole in the ice</td>
</tr>
</tbody>
</table>
Challenges

● Reference materials may be developed for non-phonological purposes.
● Orthography may obscure phonology or phonetics.
● Phonological and morphological analyses presented in grammars may be more surface-oriented (in materials developed for indigenous communities for ex.) vs. more abstract and depth-oriented, and this may lead to differing analytical results concerning prosody and edge effects.
● Solution (in-progress): utilize audio resources to transcribe in IPA and investigate suprasegmental prosody…
Audio-based resources

- **Talking dictionaries** (e.g., [https://itwewina.altlab.app](https://itwewina.altlab.app) [Arppe et al. 2022]; Fisher, Leman, Pine, and Sanchez 2006; Genee & Junker 2018; Ojibwe People’s Dictionary 2022)

- **Speech corpora** (e.g., Spoken Dictionary of Maskwacîs Cree, cf. [https://speech-db.altlab.app/maskwacis/entries](https://speech-db.altlab.app/maskwacis/entries) [Arppe et al. 2022])
  - Create transcriptions in the International Phonetic Alphabet (IPA)
  - Create phonological generalizations based on transcriptions
  - Possibility to study other phenomena like stress

- **Working with speakers to record datasets based on orthography**
References


https://doi.org/10.1093/oso/9780198778264.003.0011

https://doi.org/10.1093/oso/9780198778264.003.0011


Weber et al. forthcoming. Blackfoot Words: A lexical database of Blackfoot legacy sources. Accepted to *Language Resources and Evaluation*.