Reconstructing Negation Morphemes and Constructions in the Tupí-Guaraní Family

Natalia Chousou-Polydouri,¹ Noé Gasparini,¹ Lev Michael,² Zachary O’Hagan,² Françoise Rose¹

∼

AMAZONICAS VI
Leticia, Amazonas, Colombia
May 24, 2016
Introduction

• Goal:
  • Describe the evolution of negation in the Tupí-Guaraní (TG) family
    • Based on ancestral state reconstruction methods from computational phylogenetics

• Results:
  • Confirm previous results concerning negation in proto-TG (PTG)
  • Identify new elements that reconstruct to intermediate groupings within the family
  • Clarify the evolution of reconstructed elements to daughter languages

• Methodological contributions
  • Innovative extension of computational phylogenetics from the lexicon to morphosyntax
  • Increase the rigor of morphosyntactic reconstruction by quantifying probability of reconstructions to particular nodes
Organization

1. TG background
2. Cognacy between morphemes and between constructions
3. Computational phylogenetic methods
4. Data collection and cognate sets development
5. Ancestral state reconstruction
6. Results for TG
7. Conclusions
Tupí-Guaraní

- Subgroup of Tupí family with > 45 languages

Figure 1: Tupí-Guaraní Classification (Michael et al. 2015)
Tupí-Guaraní & Morphosyntax

- Sample of comparative work in morphosyntax

- Less morphosyntactic reconstruction

- Even less morphosyntax used in classification
  - Rodrigues (1984/1985); Dietrich (1990); Mello (2000); Rodrigues & Cabral (2002); Michael et al. (2015)

- Additional comparative work on Tupí family as a whole
Negation in Tupí-Guaraní

  - *nda...-i  
    STANDARD NEGATION
  - *eʔim  
    NOMINAL PRIVATIVE, DEPENDENT VERBS
  - *ruã/*ruĩ  
    CONSTITUENT NEGATION
  - *eme  
    NEGATIVE IMPERATIVE
  - *ani  
    ‘no’
  - *-ts(o)wε  
    (additionally required with some TAM)

- In this previous work:
  - No explicit reconstruction of morphemes at intermediate nodes
  - No explicit reconstruction of negation constructions

- In this talk, we provide explicit reconstructions of the presence of morphemes and constructions at intermediates nodes
  - using phylogenetic methods applied to both morpheme and construction cognate sets
Cognacy: Morphemes and Constructions

• What do we mean when we say that two constructions are cognate?
• First think about what we mean when we say that morphemes are cognate:
  • Morpheme: pairing of phonological form and meaning/function
  • Morphemes in different languages are cognate if they descend from a common ancestor
    • phonological form and meaning/function may each change with respect the ancestral form
• Constructions:
  • Clustering of: 1) form of elements in construction; 2) syntagmatic relationship between elements; and 3) meaning/function of the set of elements
  • Construction are similarly cognate if the descend from a common ancestor, although the form of the elements, their syntagmatic relationship, and the meaning/function of set of elements may change
Construction Cognacy Example

- Negative deontic
  - \( t\text{é}=i\text{mé} \ V \) TUPINAMBÁ
- Standard negation
  - \( t\text{íma} \ V \) KOKAMA
- These two constructions descend from the same ancestral construction (a construction essentially identical to the Tupinambá one), and so are cognate
  - The elements that compose these constructions descend from the same ancestral set of elements, but have changed form in Kokama
  - The syntagmatic configuration of the elements has likewise changed in Kokama (becoming a single element)
- The meaning/function of the construction has changed, so they are not synchronically the same construction
Morpheme Cognacy Example

• Tupinambá =i-même occurs in another construction
  • V=i-même
  • Reflex of PTG *eme (Jensen 1998a:549)
  • Reflexes of *eme are present in numerous languages of different branches (e.g., Kamaiurá -em)
  • In Kokama, tìma is the only trace of a reflex of *eme
    • Analogous to Latin vir ‘man’ and English werewolf
  • Thus tìma is included in a set with =i-même and -em as an instance of morpheme cognacy
Computational Phylogenetic Methods

• Phylogenetic methods find evolutionary trees that account for distributions of traits across languages
• Traits are modeled as binary or multistate characters
  • Binary characters (0, 1, ?) reflect the presence or absence of cognates
  • Multistate characters (0, 1, 2, . . . , ?) reflect which of possibly many functions a construction expresses
• The algorithm searches a space of possible trees
• Tree search is guided by parsimony optimization
  • The analysis attempts to minimize the number of necessary independent innovations (maximum likelihood)
• The resulting trees both identify subgroupings and track evolution of traits
  → phylogenetic trees distinguish innovations from retentions
Data Set

- Our data set consists of constructions that express the following 11 negation-related functions in 27 TG languages and two non-TG Tupian languages:
  - standard negation
  - focused constituent negation
  - negative imperative
  - negative directive
  - denominal privative
  - deverbal privative
  - free negative answer ‘no’
  - free prohibitive ‘don’t’
  - metalinguistic negation
  - dedicated intentional/desiderative negation
  - dedicated future negation
Building Cognate Sets

- Morpheme cognate sets
  - 45 sets
  - 20 sets with at least two languages
- Construction cognate sets
  - 23 sets with at least two languages and distinct from morpheme cognate sets
**eʔim** Morpheme Cognate Set

<table>
<thead>
<tr>
<th>SOU</th>
<th>TUE</th>
<th>FRM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paraguaya</td>
<td>Negative Directive</td>
<td>&lt;t(V)-PR-V-ʔ (PRO)&gt;</td>
</tr>
<tr>
<td>Kalowa</td>
<td>Privative (Denominal)</td>
<td>&lt;N-ʔi&gt;</td>
</tr>
<tr>
<td>Kalowa</td>
<td>Negative Directive</td>
<td>&lt;t-PR-V-ʔ (PRO)&gt;</td>
</tr>
<tr>
<td>Kalowa</td>
<td>Privative (Denominal)</td>
<td>&lt;N-ʔ&gt;</td>
</tr>
<tr>
<td>Mbyá</td>
<td>Privative (Denominal)</td>
<td>&lt;N-ʔ&gt;</td>
</tr>
<tr>
<td>Omagua</td>
<td>Privative (Denominal)</td>
<td>/N-ima/</td>
</tr>
<tr>
<td>Kokama</td>
<td>Privative (Denominal)</td>
<td>/N-ima/</td>
</tr>
<tr>
<td>Tupinambá</td>
<td>Main Clause Negation</td>
<td>&lt;PR-V-ʔ&gt;</td>
</tr>
<tr>
<td>Tupinambá</td>
<td>Privative (Denominal)</td>
<td>&lt;N-ʔ&gt;</td>
</tr>
<tr>
<td>Tepirape</td>
<td>Focused Constituent Negation</td>
<td>&lt;N, V, ADV, POSTP-eʔim&gt;</td>
</tr>
<tr>
<td>Tepirape</td>
<td>Privative (Denominal)</td>
<td>&lt;N-eʔim&gt;</td>
</tr>
<tr>
<td>TocantinsA...</td>
<td>Focused Constituent Negation</td>
<td>&lt;N-iʔim&gt;</td>
</tr>
<tr>
<td>Parakanã</td>
<td>Main Clause Negation</td>
<td>&lt;na-ne-n=PR-V-iʔim&gt;</td>
</tr>
<tr>
<td>Parakanã</td>
<td>Focused Constituent Negation</td>
<td>&lt;N-iʔim&gt;</td>
</tr>
<tr>
<td>Araweté</td>
<td>Privative (Denominal)</td>
<td>&lt;N-imeʔe&gt;</td>
</tr>
<tr>
<td>Parintintin</td>
<td>Privative (Denominal)</td>
<td>&lt;N-eʔim-REL&gt;</td>
</tr>
<tr>
<td>Kamaurá</td>
<td>Focused Constituent Negation</td>
<td>&lt;N, POSTP, ADV, V-ʔim&gt;</td>
</tr>
<tr>
<td>Kamaurá</td>
<td>Privative (Denominal)</td>
<td>&lt;N-eʔim&gt;</td>
</tr>
<tr>
<td>Wayampi</td>
<td>Negative Directive</td>
<td>/t-PR-V,N-ʔi/</td>
</tr>
<tr>
<td>Wayampi</td>
<td>Negative Directive</td>
<td>/sa-ʔi/</td>
</tr>
<tr>
<td>Wayampi</td>
<td>Negative Directive</td>
<td>/si-ʔi/</td>
</tr>
<tr>
<td>Wayampi</td>
<td>Privative (Denominal)</td>
<td>/sane-V,N-ʔi/</td>
</tr>
<tr>
<td>Wayampi</td>
<td>Main Clause Negation</td>
<td>//PR-ʔim/</td>
</tr>
</tbody>
</table>
Building Cognate Sets

- Morpheme cognate sets
  - 45 sets
  - 20 sets with at least two languages
- Construction cognate sets
  - 23 sets with at least two languages and distinct from morpheme cognate sets
### Working on cognate sets

#### NEG_PRIV: N-e'îm (const) [AL]

<table>
<thead>
<tr>
<th>SOU</th>
<th>TUE</th>
<th>FRM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kalowá</td>
<td>Privative (Denominal)</td>
<td>&lt;N-e'îm&gt;</td>
</tr>
<tr>
<td>Kalowá</td>
<td>Privative (Denominal)</td>
<td>&lt;N-e'yî&gt;</td>
</tr>
<tr>
<td>Mbyá</td>
<td>Privative (Denominal)</td>
<td>&lt;N-e'yî&gt;</td>
</tr>
<tr>
<td>Omagua</td>
<td>Privative (Denominal)</td>
<td>/N-ima/</td>
</tr>
<tr>
<td>Kokama</td>
<td>Privative (Denominal)</td>
<td>/N-ima/</td>
</tr>
<tr>
<td>Tupinambá</td>
<td>Privative (Denominal)</td>
<td>&lt;N-eym&gt;</td>
</tr>
<tr>
<td>Tapirapé</td>
<td>Privative (Denominal)</td>
<td>&lt;N-e'yîm&gt;</td>
</tr>
<tr>
<td>TocantinsA...</td>
<td>Focused Constituent Negation</td>
<td>&lt;N-i'îm&gt;</td>
</tr>
<tr>
<td>Parakanã</td>
<td>Focused Constituent Negation</td>
<td>&lt;N-i'îm&gt;</td>
</tr>
<tr>
<td>Araweté</td>
<td>Privative (Denominal)</td>
<td>&lt;N-ime?e&gt;</td>
</tr>
<tr>
<td>Parintintin</td>
<td>Privative (Denominal)</td>
<td>&lt;N-e'yîm-REL&gt;</td>
</tr>
<tr>
<td>Kamaiurá</td>
<td>Privative (Denominal)</td>
<td>&lt;N-e'yîm&gt;</td>
</tr>
<tr>
<td>Wayampí</td>
<td>Privative (Denominal)</td>
<td>/N-ý/</td>
</tr>
<tr>
<td>Ka'apor</td>
<td>Privative (Denominal)</td>
<td>&lt;N ym&gt;</td>
</tr>
</tbody>
</table>
Tracking Morpheme and Construction Evolution with Ancestral State Reconstruction

- Ancestral state reconstructions:
  - Gives the probability of presence of given trait at a given non-terminal node in a tree
  - Traces evolution of a feature
- Tree on which traits evolve is presupposed, i.e. has to come from some other analysis, e.g. lexical phylogenetic analysis (Michael et al. 2015)
- Morpheme cognate evolution:
  - Maximum likelihood reconstruction
  - Model assumes that gains are much less frequent than losses ($\sim 1:15$ Bayesian estimate)
- Construction cognate evolution:
  - Parsimony reconstruction
  - Model is a state network
Examples of Morpheme Ancestral State Reconstructions

- Presence of reflex of *ani
- Presence of reflex of *e?am
Presence of Reflex of *ani
Presence of Reflex of *eʔam
PTG Negation Morphemes

✓ *eʔim (PTG + Awetí)
✓ *ruã/*ruĩ
✓ *eme
✓ *ani (PTG + Awetí)
X *ts(o)w e
Examples of Construction Ancestral State Reconstructions

- Presence of reflex of \(*PR_{IMP}^{\text{-}}\text{V} \ eme\) construction (negative imperative)
  - Reconstructs to PTG
- Presence of reflex of \(*ani\ PR-V\) (negative directive or negative imperative)
  - Innovation in Guaranian subgroup
Presence of $^{*}$PR_{IMP-V} eme Construction
Presence of *ani PR-V
PTG Negation Constructions

<table>
<thead>
<tr>
<th>Construct</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>*PR-V eme</td>
<td>NEGATIVE IMPERATIVE</td>
</tr>
<tr>
<td>*ta=PR-V eme</td>
<td>NEGATIVE DIRECTIVE</td>
</tr>
<tr>
<td>*nda=PR-V-i</td>
<td>STANDARD NEGATION</td>
</tr>
<tr>
<td>*N-eʔiʔm</td>
<td>DENOMINAL PRIVATIVE</td>
</tr>
<tr>
<td>*ani</td>
<td>‘no’ (free negation)</td>
</tr>
</tbody>
</table>
Reconstruction of Functions

• There are potentially a large number of functions associated with the reflexes of a given ancestral construction
  • As such, function characters are not binary, but multistate \( (0, 1, 2, \ldots) \)
• Generally speaking, transitions between these multiple states are constrained, in the sense that not all states can be directly reached from a particularly given state, or not reached with equal ease
  • The restrictions on these transitions are modeled as state network (see next slide)
• We now examine the evolution of the element \(^*emē\), which participates in a variety of constructions
State Network for Function of *eme

Negative Directive ← Negative Intentional

Standard Negation
Focused Constituent Negation

Negative Imperative ← Negative Directive
Negative Imperative ↓ Negative Directive

3 3 4 3
Evolution of functions of *eme

Tupinambá 'Lemos Barbosa 1956:91-92'

*e-*pysyk umé
2sg.imp-3.P-take neg.imp
'Don’t take it.'

*t’ o-*pysyk umé
dir 3A-3P-take neg.dir
'He should not take it!'

‘He should not take it!’
Conclusions

- 6 PTG negators out of 45 cognate sets
  - *ani
  - *eʔim
  - *nda
  - -*i
  - *ruã/*ruĩ
  - *eme
- 5 PTG negative constructions
- Many reconstructions
- Some examples of grammaticalization (e.g., *eʔam)
- Many instances of functional extension (e.g., -ã in Siriono)
Conclusions

• Phylogenetic methods and morphosyntax
  • Morphology already treated just like lexicon
  • Syntactic data can also be used
  • Both add information refining the classification

• Ancestral state reconstruction techniques help
  • Determining where in the tree an element can be reconstructed
  • Determining the proto-function of that element
Acknowledgments

- Data harvesters
  - Keith Bartolomei, Natalia Chousou-Polydouri, Emily Clem, Windy Daviet, Noé Gasparini, Paola Granado Columba, Magdalena Lemus Serrano, Lev Michael, Zachary O’Hagan, Françoise Rose

- Additional data contributors
  - Eva-Maria Rößler, Sérgio Meira

- RefLex development and support
  - Sébastien Flavier

- Funding
  - France-Berkeley Fund, LabEx ABLAN
Lexicostatistics ≠ Phylogenetics

- Lexicostatistical Methods (e.g. NeighborNet, SplitsTree)
  - Lexicostatistical methods do not evaluate evolutionary trees
  - They instead compute a single number – % of shared cognates – for each pair of languages
  - Languages are then clustered on the basis of overall similarity, conflating shared innovations and shared retentions
- Phylogenetic Methods
  - All cognate sets are evaluated individually, and the specific information they bear for subgrouping is preserved
  - Thousands of trees are individually evaluated by optimizing all characters on each one
  - Only shared innovations are considered for subgrouping
  - As a result, phylogenetic methods are less likely to be “fooled” by borrowing than lexicostatistical methods


DERBYSHIRE, DESMOND C. 1994. Clause Subordination and


MICHAEL, LEV; NATALIA CHOOUSSO-POLYDOURI; KEITH BARTOLOMEI; ERIN DONNELLY; VIVIAN WAUTERS; SÉRGIO MEIRA; and ZACHARY O’HAGAN. 2015. A Bayesian Phylogenetic Classification of Tupí-Guaraní. *LIAMES* 15.193–221.


