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 Rose[1]

$\sim$

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

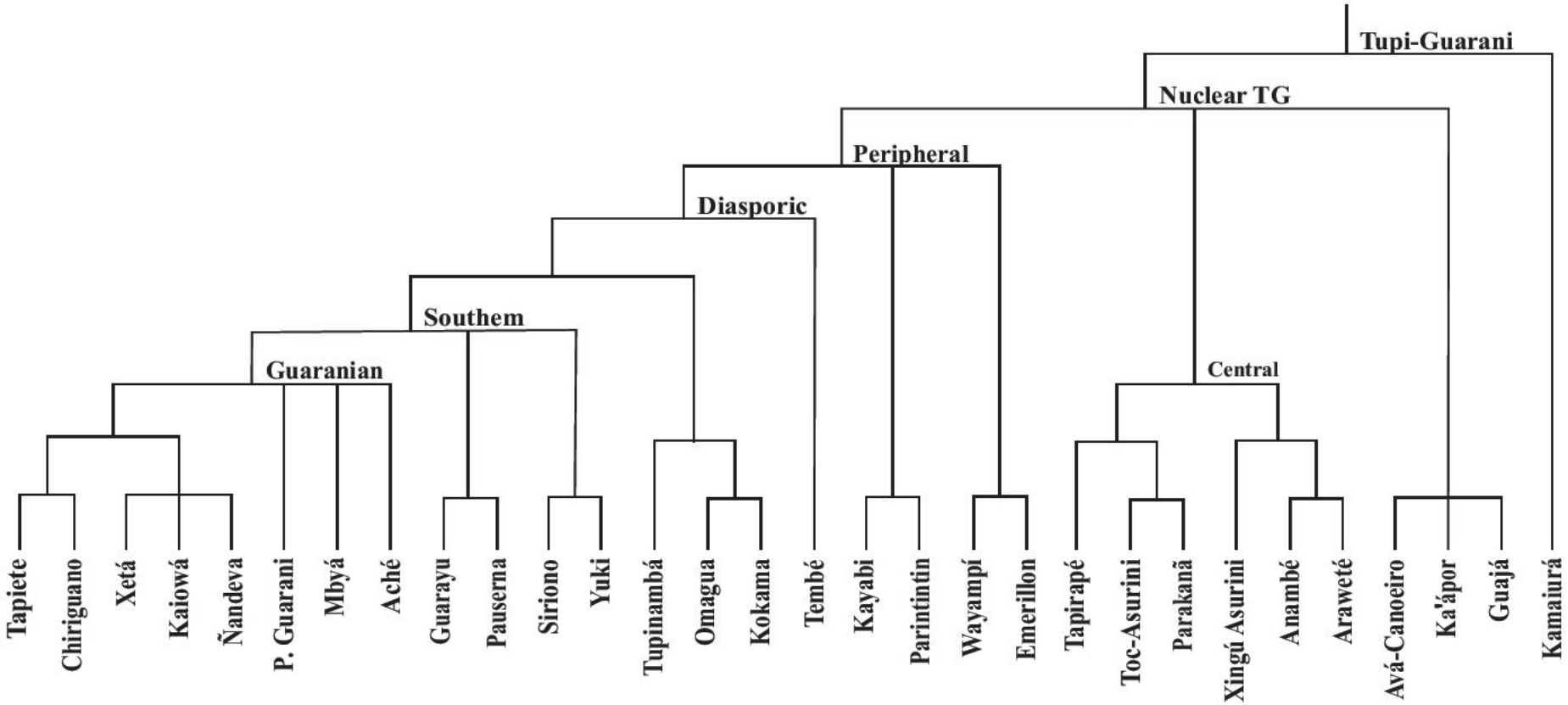


Figure 1: Tupí-Guaraní Classification (Michael et al. 2015)

Tupí-Guaraní & Morphosyntax

- Sample of comparative work in morphosyntax
  - Cabral (2001, 2007); Cabral & Rodrigues (2005); Derbyshire (1994); Dietrich (1977, 1994, 2000, 2001, 2006, 2010); Edelweiss (1969–1972); Gasparini (2011); Jensen (1998b); Payne (1994); Queixalós (2006); Rößler (2014); Rose (2015); Sales (2002, 2007); Vieira (2007, 2014)
- Less morphosyntactic reconstruction
  - Dietrich (2009); Jensen (1987, 1990, 1997, 1998a, 1999); Meira & Drude (2013); Schleicher (1998)
- 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

- Reconstructions of PTG negation morphemes (Jensen (1998a:545-549) & Schleicher (1998:298-300)):
  - **nda...-i* STANDARD NEGATION
  - **eʔɨm* NOMINAL PRIVATIVE, DEPENDENT VERBS
  - **ruã/*ruĩ* CONSTITUENT NEGATION
  - **eme* NEGATIVE IMPERATIVE
  - **ani* 'no'
  - **-ts(o)^{w}e* (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
  - *té=ɨmé* V TUPINAMBÁ
- Standard negation
  - *tɨma* 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 =*ɨmé* occurs in another construction
  - V=*ɨmé* NEG IMPERATIVE
- Reflex of PTG **eme* (Jensen 1998a:549)
- Reflexes of **eme* are present in numerous languages of different branches (e.g., Kamaiurá -*em*)
- In Kokama, *tɨma* is the only trace of a reflex of **eme*
  - Analagous to Latin *vir* 'man' and English *werewolf*
- Thus *tɨma* is included in a set with =*ɨmé* 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

*eʔɨm Morpheme Cognate Set

Working on cognate sets Neg: eʔɨm (morph) [AL]

27 Record(s)

SOU	TUE	FRM
Paraguaya...	Negative Directive	<t(V)-PR-V-'ỹ (PRO)>
Kaiowá	Privative (Denominal)	<N-eʔĩ>
Kaiowá	Negative Directive	<t-PR-V-'ỹ (PRO)>
Kaiowá	Privative (Denominal)	<N-e'ỹ>
Mbyá	Privative (Denominal)	<N-e'ỹ>
Omagua	Privative (Denominal)	/N-ima/
Kokama	Privative (Denominal)	/N-ima/
Tupinambá	Main Clause Negation	<PR-V-eym>
Tupinambá	Privative (Denominal)	<N-eym>
Tapirapé	Focused Constituent Negation	<N, V, ADV, POSTP-e'ym>
Tapirapé	Privative (Denominal)	<N-e'ym>
TocantinsA...	Focused Constituent Negation	<N-iʔim>
Parakanã	Main Clause Negation	<na~ne~n=PR-V-iʔím>
Parakanã	Focused Constituent Negation	<N-iʔím>
Arawetè	Privative (Denominal)	<N-imeʔe>
Parintintin	Privative (Denominal)	<N-e'ym-REL>
Kamaiurá	Focused Constituent Negation	<N, POSTP, ADV, V-e'ym>
Kamaiurá	Privative (Denominal)	<N-ey'm>
Wayampi	Negative Directive	/t-PR-V,N-'ỹ/
Wayampí	Negative Directive	/sa-V-'ỹ/
Wayampí		/si-V-'ỹ/
Wayampí		/sane-V,N-'ỹ/
Wayampí	Privative (Denominal)	/N-'ỹ/
Ka'ápor	Main Clause Negation	/(PR)-V ʔim/

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

*N-*eʔɨm* Construction Cognate Set

Working on cognate sets NEG_PRIV: N-e'ɨm (const) [AL]

14 Record(s)

	SOU	TUE	FRM
✓	Kaiowá	Privative (Denominal)	<N-eʔĩ>
✓	Kaiowá	Privative (Denominal)	<N-e'ỹ>
✓	Mbyá	Privative (Denominal)	<N-e'ỹ>
✓	Omagua	Privative (Denominal)	/N-ima/
✓	Kokama	Privative (Denominal)	/N-ima/
✓	Tupinambá	Privative (Denominal)	<N-eym>
✓	Tapirapé	Privative (Denominal)	<N-e'ym>
✓	TocantinsA...	Focused Constituent Negation	<N-iʔim>
✓	Parakanã	Focused Constituent Negation	<N-iʔím>
✓	Araweté	Privative (Denominal)	<N-imeʔe>
✓	Parintintin	Privative (Denominal)	<N-e'ym-REL>
✓	Kamaiurá	Privative (Denominal)	<N-ey'm>
✓	Wayampí	Privative (Denominal)	/N-'ỹ/
✓	Ka'ápor	Privative (Denominal)	<N ym>

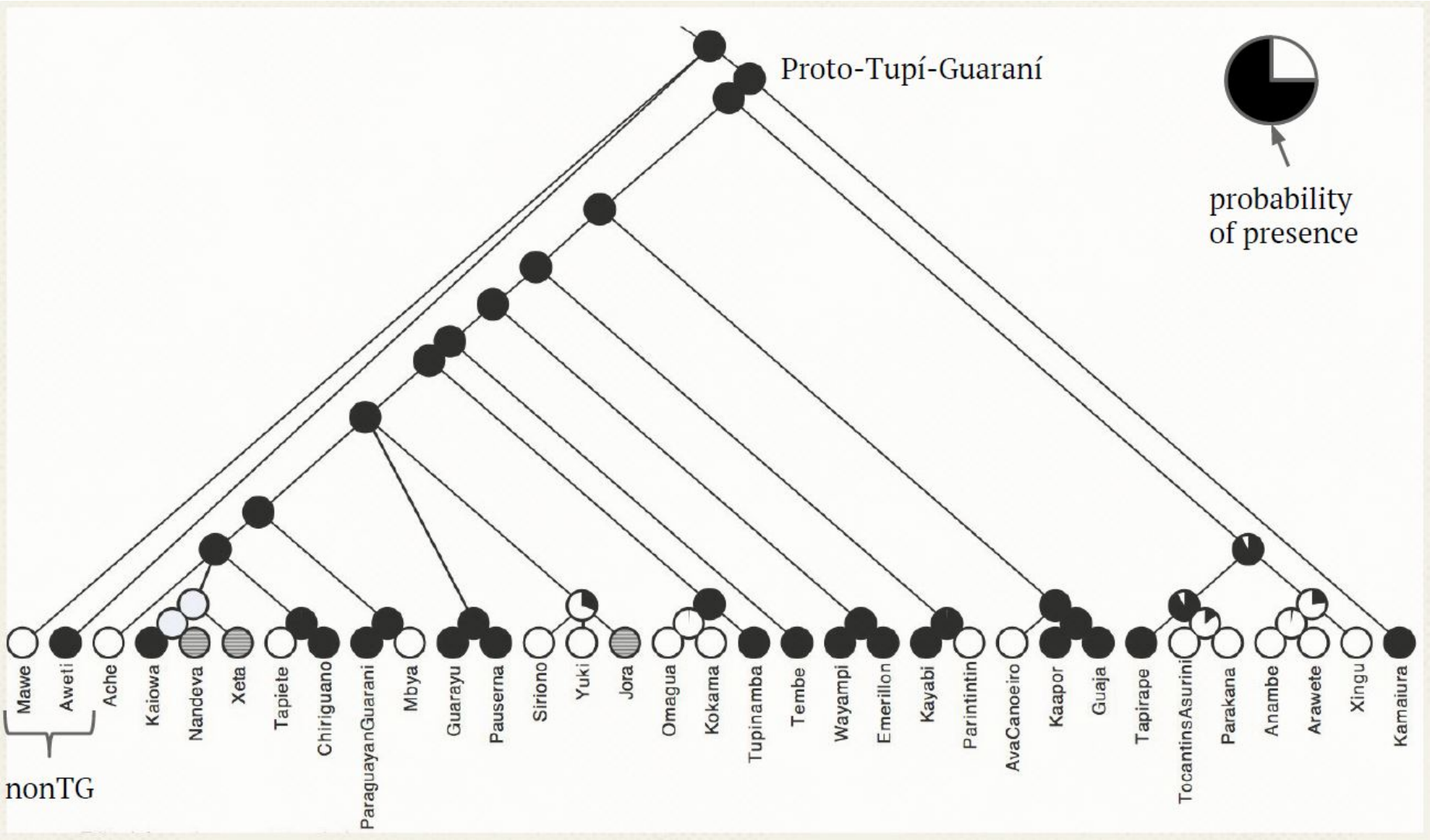
Tracking Morpheme and Construction Evolution with Ancestral State Reconstruction

- 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

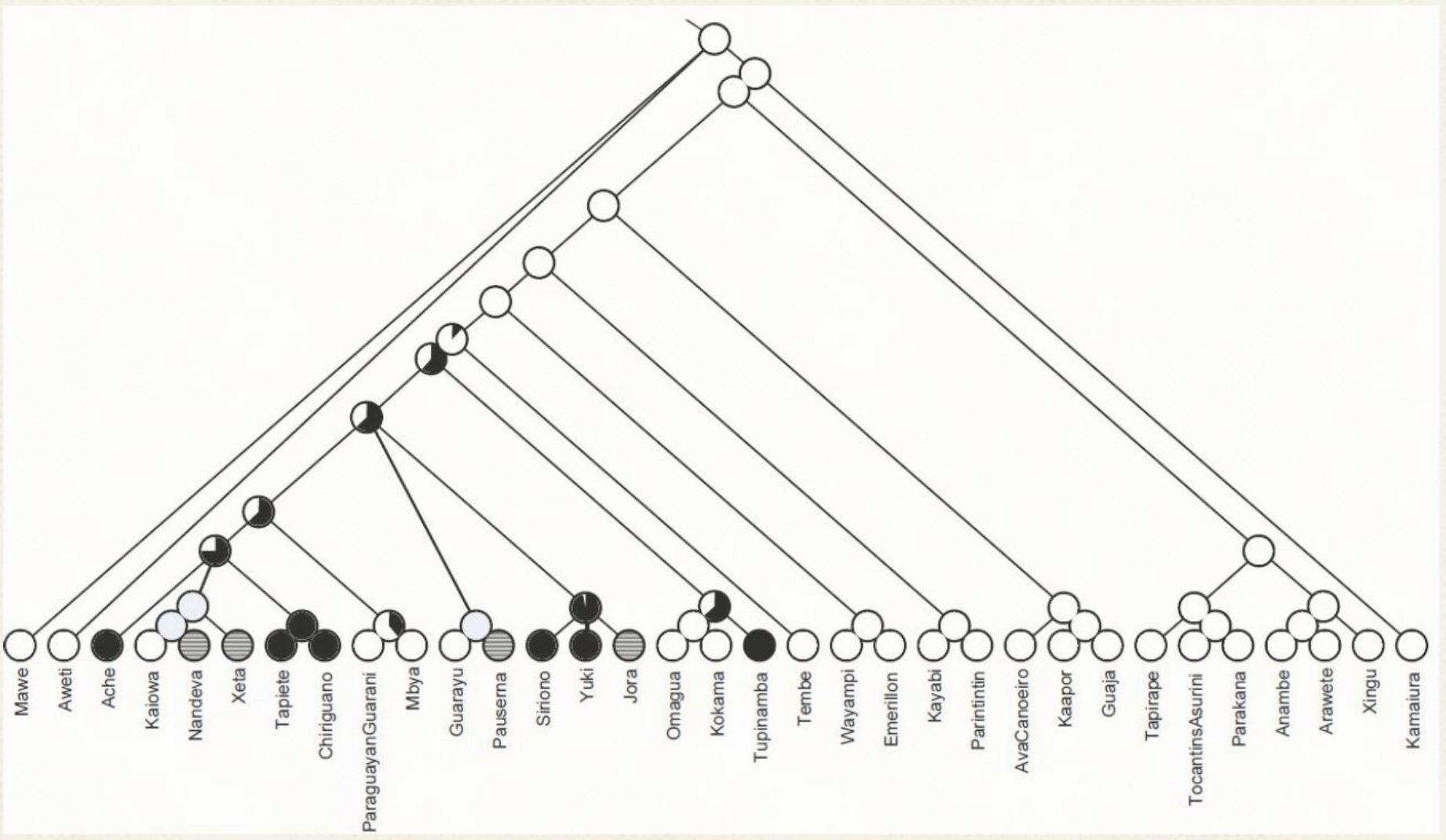
Examples of Morpheme Ancestral State Reconstructions

- Presence of reflex of **ani*
- Presence of reflex of **eʔam*

Presence of Reflex of *ani



Presence of Reflex of *eʔam



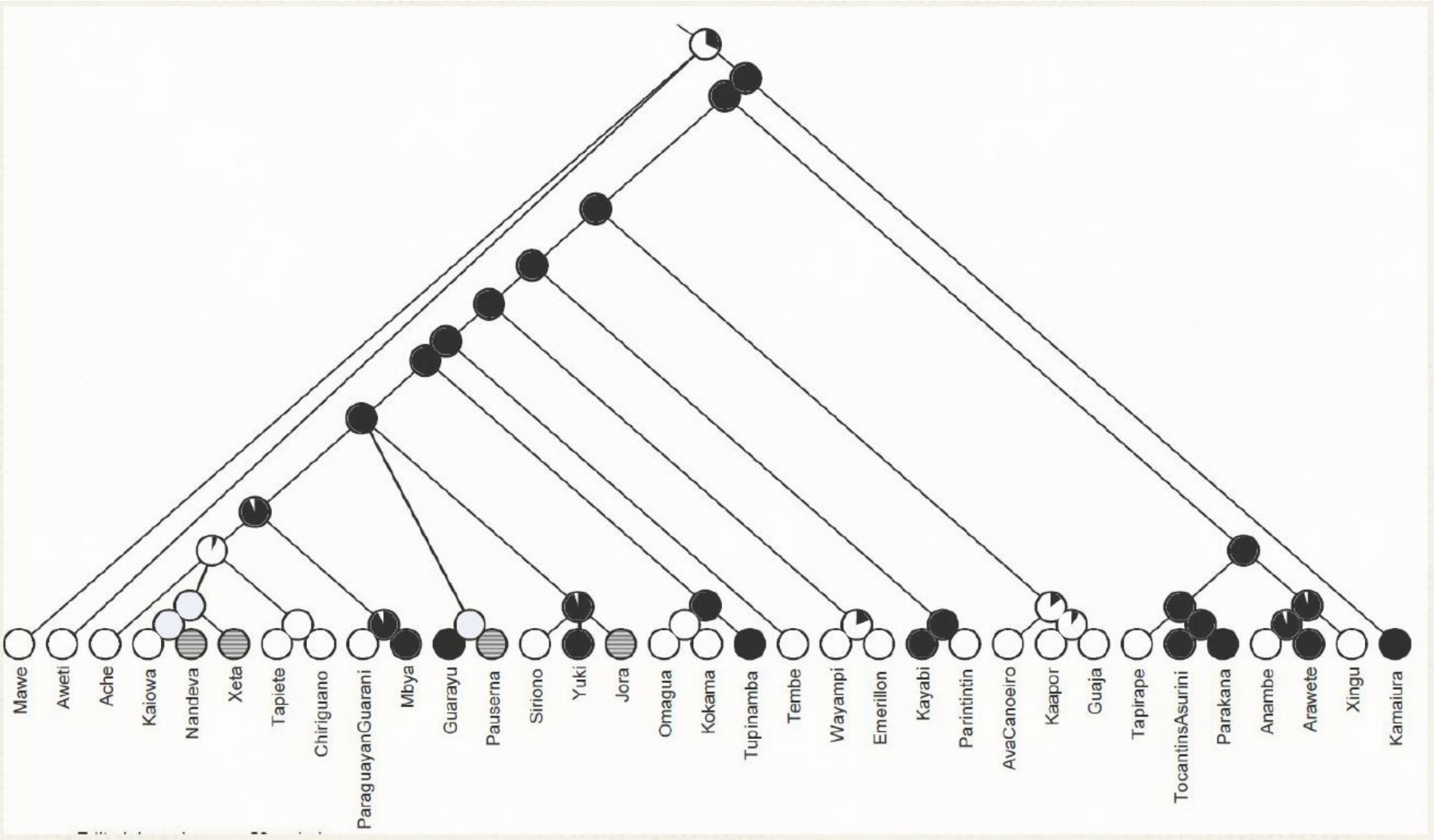
PTG Negation Morphemes

- ✓ **eʔɨm* (PTG + Awetí)
- ✓ **ruã/*ruĩ*
- ✓ **eme*
- ✓ **ani* (PTG + Awetí)
- X **ts(o)^{w}e*

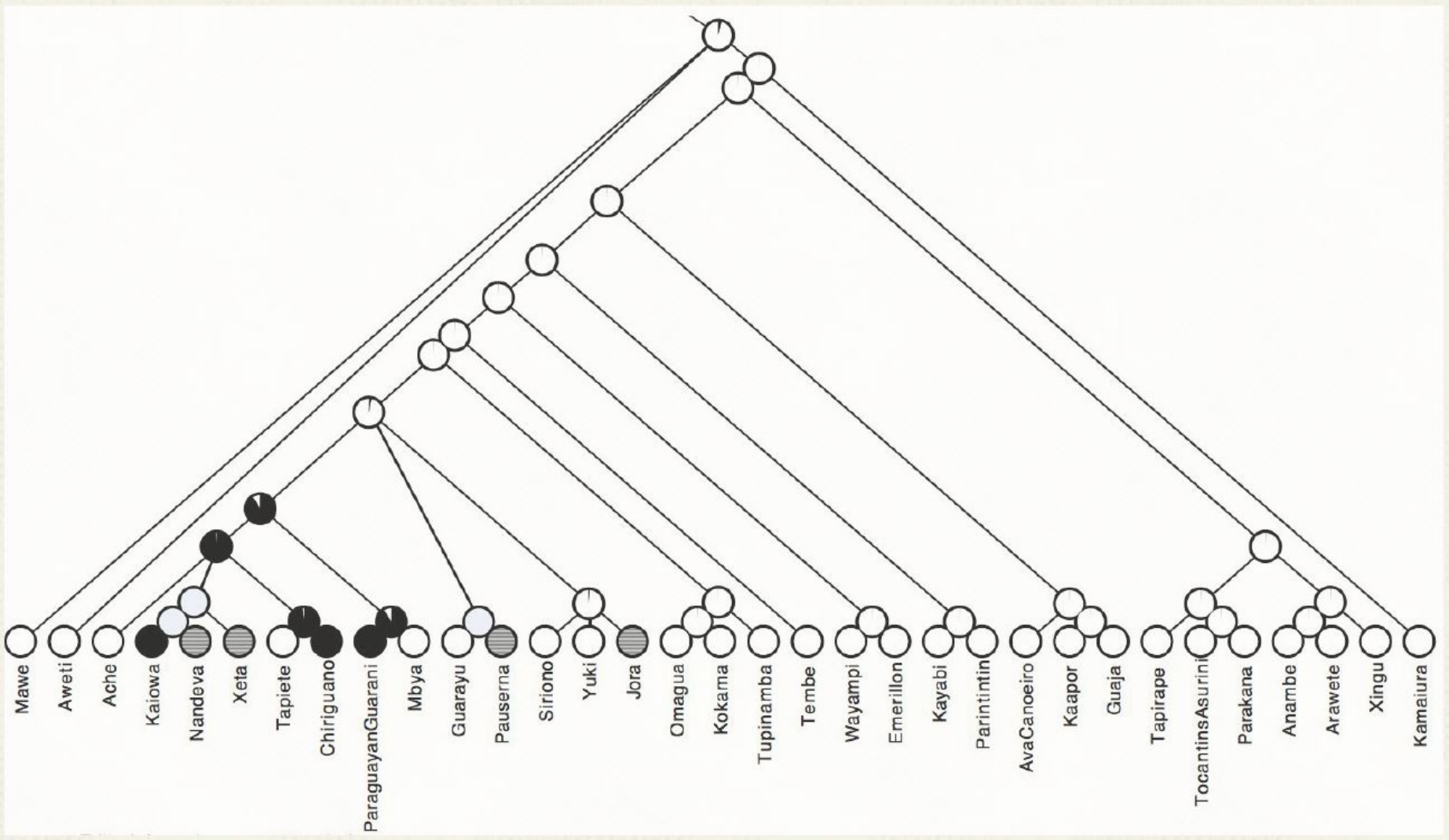
Examples of Construction Ancestral State Reconstructions

- Presence of reflex of *PR_{IMP}-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



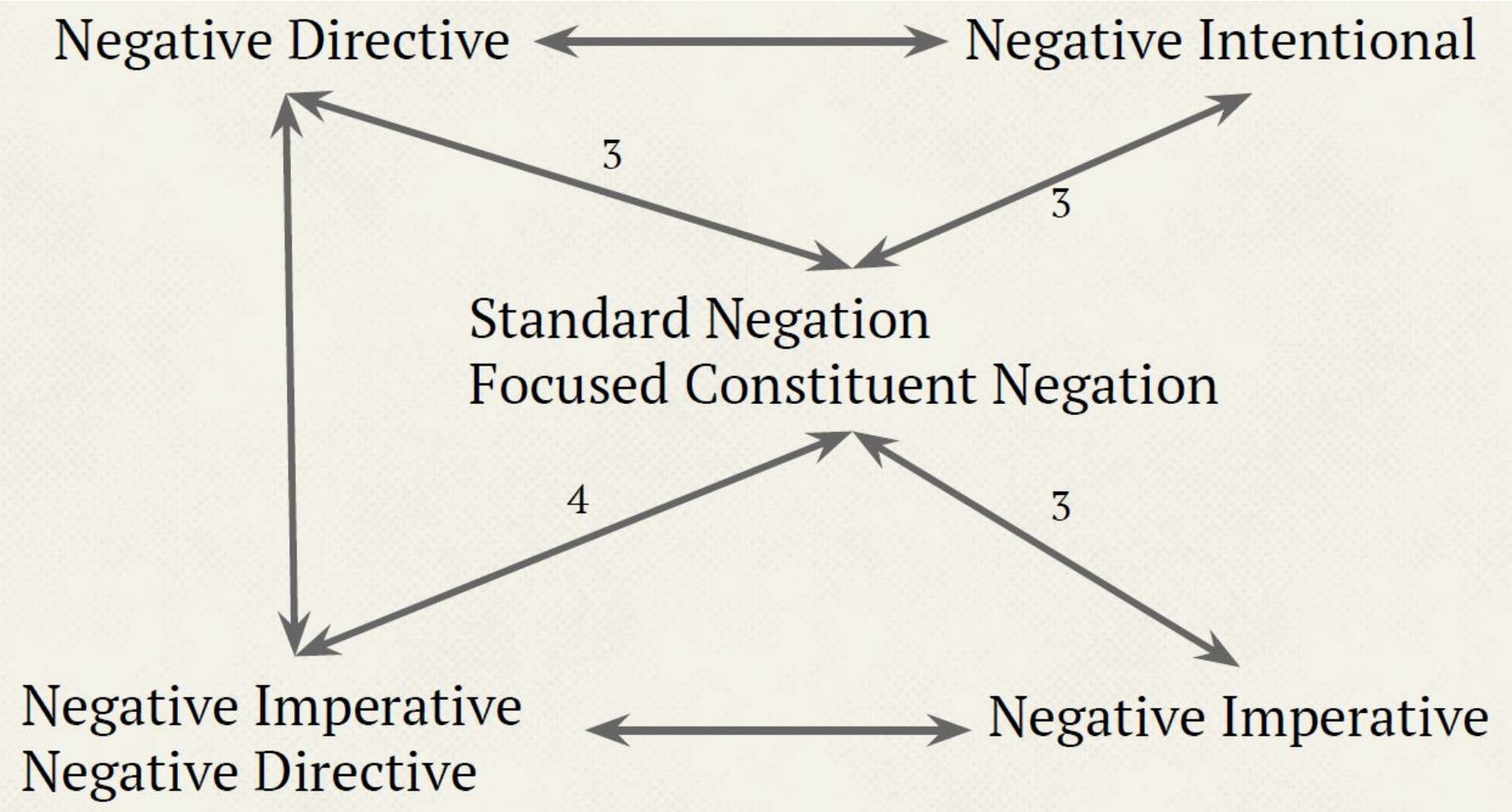
PTG Negation Constructions

✓	*PR-V *eme*	NEGATIVE IMPERATIVE
?	**ta*=PR-V *eme*	NEGATIVE DIRECTIVE
✓	**nda*=PR-V-*i*	STANDARD NEGATION
✓	*N-*eʔɨm*	DENOMINAL PRIVATIVE
✓	**ani*	‘no’ (free negation)

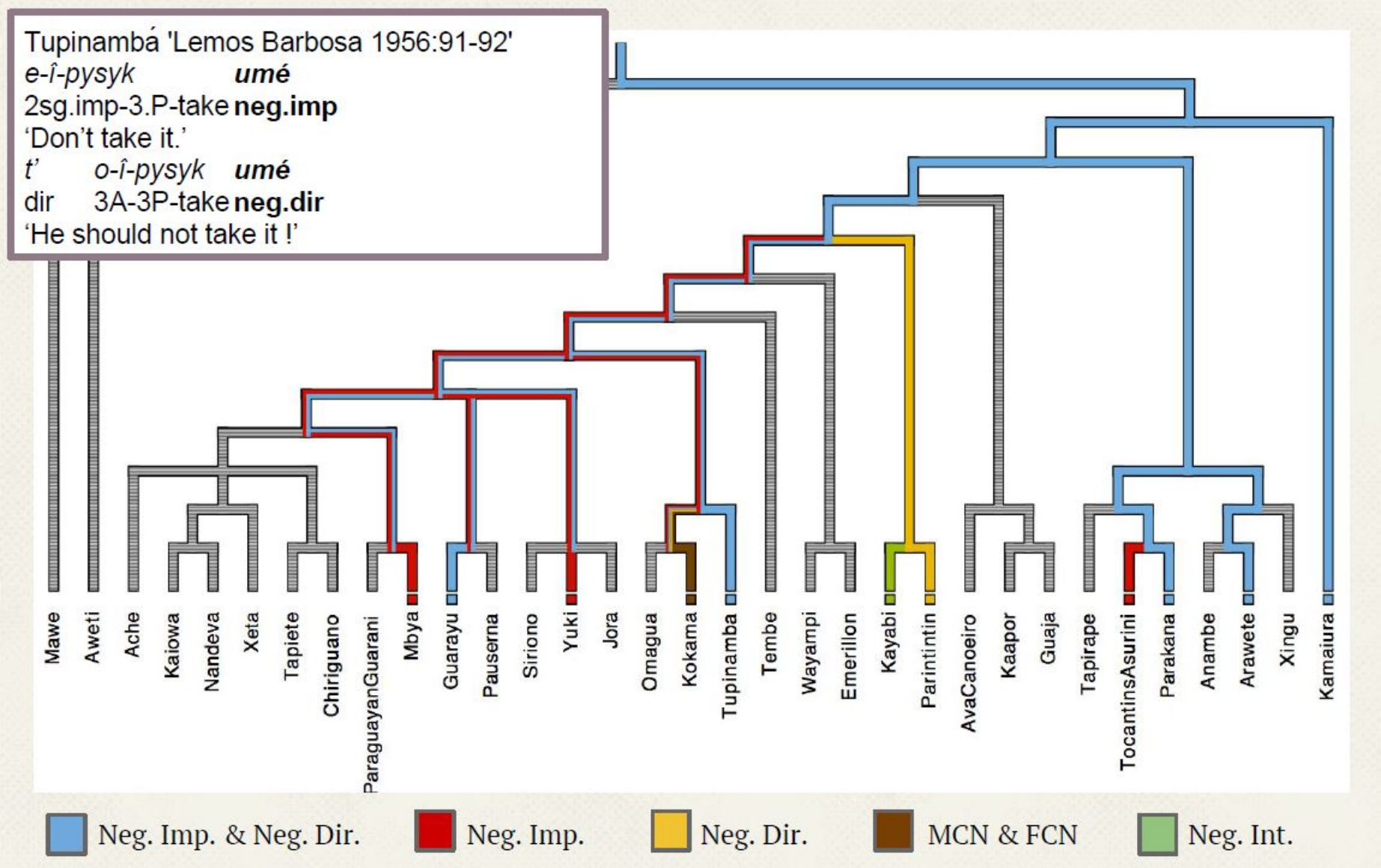
Reconstruction of Functions

- 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, . . .)
- 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 **eme*, which participates in a variety of constructions

State Network for Function of **eme*



Evolution of functions of *eme



Conclusions

- 6 PTG negators out of 45 cognate sets
  - **ani*
  - **eʔɨm*
  - **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
  - Sébastian Flavier
- Funding
  - France-Berkeley Fund, LabEx ABLAN

Lexicostatistics $\neq$ 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

CABRAL, ANA SUELLY ARRUDA CÂMARA. 2001. Observações sobre a história do morfema -a da família tupí-guaraní. *Des noms et des verbes en tupi-guarani: Etat de la question*, ed. by Francisco Queixalós, 133–162. Munich: LINCOLM Europa.

CABRAL, ANA SUELLY ARRUDA CÂMARA. 2007. L'Expression des notions de l'épistémique et de l'aléthique dans la famille tupi-guarani. *L'Énonciation médiatisée II. Le traitement épistémologique de l'information: Illustrattions amérindiennes et caucasiennes*, ed. by Zlatka Guentchéva and Jon Landaburu, no. 63 in Bibliothèque de l'Information Grammaticale, 267–292. Louvain/Paris: Éditions Peeters.

CABRAL, ANA SUELLY ARRUDA CÂMARA, and ARYON DALL'IGNA RODRIGUES. 2005. O desenvolvimento do gerúndio e do subjuntivo em tupí-guaraní. *Novos estudos sobre línguas indígenas*, ed. by Aryon Dall'Igna Rodrigues and Ana Suelly Arruda Câmara Cabral, 47–58. Brasília: Editora, Universidade de Brasília.

DERBYSHIRE, DESMOND C. 1994. Clause Subordination and

Nominalization in Tupi-Guaranian and Cariban Languages. *Revista Latinoamericana de Estudios Etnolingüísticos* 8.179–198.

DIETRICH, WOLF. 1977. Las categorías verbales (partes de la oración) en tupí-guaraní. *Indiana* 4.245–261.

DIETRICH, WOLF. 1990. *More Evidence for an Internal Classification of Tupí-Guaraní Languages*. Berlin: Gebr. Mann Verlag.

DIETRICH, WOLF. 1994. Word Formation, Syntax or Noun Classification? Tupí-Guaraní *mba'e*: Between Lexicon and Grammar. *Revista Latinoamericana de Estudios Etnolingüísticos* 8.110–124.

DIETRICH, WOLF. 2000. Problema de la categoría del adjetivo en las lenguas tupí-guaraníes. *Indigenous languages of latin america (illa)*, ed. by Hein van Der Voort and Simon van De Kerke, *CNWS Publications*, vol. 1, 255–263. Leiden: Research School CNWS, Leiden University.

DIETRICH, WOLF. 2001. Categorias Lexicais nas Línguas Tupí-Guaraní (VisaÌČo Comparativa). *Des noms et des verbes en tupi-guarani: etat de la question*, ed. by Francisco Queixalós, 23–37. Munich: LINCOLM Europa.

DIETRICH, WOLF. 2006. La categoria del "frustrativo" en las lenguas tupí-guaraníes. *Guaraní y mawetí-tupí-guaraní: Estudos históricos y descriptivos sobre una familia lingüística de américa del sur*, ed. by Wolf Dietrich and Haralambos Symeonidis, 71–85. Berlin: LIT Verlag.

DIETRICH, WOLF. 2009. Cambio del orden de palabras en lenguas tupí-guaraníes. *Cadernos de Etnolingüística* 1.1–11.

DIETRICH, WOLF. 2010. O tronco tupi e as suas famílias de línguas: ClassificaçaÌČo e esboço tipológico. *O português e o tupi no brasil*, ed. by Wolf Dietrich and Volker Noll, 9–26. SaÌČo Paulo: Editora Contexto.

EDELWEISS, FREDERICO G. 1969–1972. *Gûasú* e *usú* na diacronia das línguas e ddialeto tupi-guaranis. *Revista do Instituto de Estudos Brasileiros* 7–12.

GASPARINI, NOÉ. 2011. *L'expression du nombre dans les langues tupi-guarani*. MA thesis, Université Lumière Lyon 2.

JENSEN, CHERYL. 1987. Object-prefix Incorporation in Proto Tupí-Guaraní Verbs. *Language Sciences* 9.45–55.

JENSEN, CHERYL. 1990. Cross-Referencing Changes in Some Tupí-Guaraní Languages. *Amazonian linguistics: Studies in lowland south american languages*, ed. by Doris L. Payne, 117–158. Austin: University of Texas Press.

JENSEN, CHERYL. 1997. *Coreferential Marking in Tupí-Guaraní Languages*. Düsseldorf: Talk at the XVIII International Conference on Historical Linguistics.

JENSEN, CHERYL. 1998a. Comparative Tupí-Guaraní Morphosyntax. *Handbook of Amazonian Languages*, ed. by Desmond C. Derbyshire and Geoffrey K. Pullum, vol. 4, 489–618. New York: Mouton de Gruyter.

JENSEN, CHERYL. 1998b. The Use of Coreferential and Reflexive Markers in Tupí-Guaraní Languages. *The Journal of Amazonian Languages* 1.1–49.

JENSEN, CHERYL. 1999. Tupí-Guaraní. *The Amazonian Languages*, ed. by R. M. W. Dixon and Alexandra Y. Aikhenvald, 125–163. Cambridge: Cambridge University Press.

MEIRA, SÉRGIO, and SEBASTIAN DRUDE. 2013. Sobre a origem histórica dos “prefixos relacionais” das línguas tupí-guaraní. *Cadernos de Etnolingüística* 5.

MELLO, ANTÔNIO AUGUSTO SOUZA. 2000. *Estudo histórico da família lingüística tupí-guaraní: Aspectos fonológicos e lexicais*. Universidade Federal de Santa Catarina PhD dissertation.

MICHAEL, LEV; NATALIA CHOUSOU-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.

PAYNE, DORIS L. 1994. The Tupí-Guaraní Inverse. *Voice, form and function*, ed. by Barbara Fox and Paul J. Hopper, 313–340. Philadelphia: John Benjamins.

QUEIXALÓS, FRANCISCO. 2006. The Primacy and Fate of Predicativity in Tupí-Guaraní. *Lexical categories and root classes in amerindian languages*, ed. by Ximena Lois and Valentina Vapnarsky, 249–287. Bern: Peter Lang.

RODRIGUES, ARYON DALL'IGNA. 1984/1985. Relações internas na família lingüística tupí-guaraní. *Revista de Antropologia* 27/28.33–53.

RODRIGUES, ARYON DALL'IGNA, and ANA SUELLY ARRUDA CÂMARA CABRAL. 2002. Revendo a classificação interna da família tupí-guaraní. *Línguas indígenas brasileiras: Fonologia, gramática e história*, ed. by Ana Suelly Arruda Câmara Cabral and Aryon Dall'Igna Rodrigues, 327–337. Belém: Editora Universitária, Universidade Federal do Pará.

RÖβLER, GASPARINI NOÉ DANIELSEN SWINTHA, EVA-MARIA. 2014. *The Southern Clade of Tupi-Guarani Languages: Adposition Issues*. Belém: Talk at AMAZONICAS V, May 29.

ROSE, FRANÇOISE. 2015. When "You" and "I" Mess Around with the Hierarchy: A Comparative Study of Tupi-Guarani Hierarchical Indexing Systems. *Boletim do Museu Paraense Emílio Goeldi* 10.347–369.

SALES, HELOÍSA MARIA MOREIRA LIMA. 2002. Subordinação em línguas da família tupí-Guaraní. *Línguas indígenas brasileiras:*

Fonologia, gramática e história, ed. by Ana Suelly Arruda Câmara Cabral and Aryon Dall'Igna Rodrigues, vol. 1, 418–424. Belém: Editora Universitária, Universidade Federal do Pará.

SALES, HELOÍSA MARIA MOREIRA LIMA. 2007. Foricidade e marcação de pessoa em línguas da família tupí-guaraní. *Línguas e culturas tupí*, ed. by Ana Suelly Arruda Câmara Cabral and Aryon Dall'Igna Rodrigues, 417–425. Campinas: Editora Curt Nimuendajú.

SCHLEICHER, CHARLES OWEN. 1998. *Comparative and Internal Reconstruction of Proto-Tupí-Guaraní*. University of Wisconsin, Madison PhD dissertation.

VIEIRA, MARCIA DAMASO. 2007. A negação sentencial em línguas da família tupí-guaraní. *Línguas e culturas tupí*, ed. by Ana Suelly Arruda Câmara Cabral and Aryon Dall'Igna Rodrigues, 277–289. Campinas: Editora Curt Nimuendajú.

VIEIRA, MARCIA DAMASO. 2014. A manifestação de tópico e foco em línguas da família tupi-guarani. *D.E.L.T.A.* 30.659–683.