

QMSS16 Presentation: Distinct Phonetic Features for Language Phylogeny

Tiago Tresoldi

May 18th, 2016



MAX-PLANCK-GESELLSCHAFT

The questions and the problems

- ▶ Is it possible (and useful) to use phonology in language phylogeny?
 - ▶ Phonology has a high rate of evolution, but borrowing is rare
 - ▶ Diachronic and synchronic data for estimation of transition probabilities
- ▶ Hypothesis: language phylogeny using distinct features (“the most basic unit of analysis in phonetic structures”), adapted from Jakobson *et al.* (1941–1956), Chomsky & Halle (1968–1983), and, in particular, Ladefoged (1971; 1993)
 - ▶ Pros: binary (and mostly non-exclusive) features (no dummy variables!), single model for vowels and consonants
 - ▶ Traditional reconstruction (such as for the IE family) uses words mostly as proxy for phonology, but we should combine lexical and phonological data
- ▶ Code & data: <https://github.com/tresoldi/qmss2016>



Table: Phonetic reflexes (snippet)

PIE	O. Irish	Latin	O. English	Lith.	Toch. B	Arm.
*p	∅	p	f	p	p	h
*ǵ ^h _i	g	h	g	ʒ	ϣ	ḫ
*k ^w	k	k ^w	h ^w	k	k	k ^h

Table: Distinct features of phonetic reflexes (snippet)

	p.ant	p.obstr	p.lab	ǵ ^h _i .obstr	ǵ ^h _i .cor
Old Irish	-	-	-	T	F
Latin	F	T	T	T	F
Old English	T	T	F	T	F
Lithuanian	F	T	T	T	T



The (very, very first) results



- ▶ Without constraints and additional data, the tree is mirroring similarities in phoneme inventories, not language history
- ▶ Ancestral State Reconstruction – do results from traditional methods match those of phylogenetic ones?
 - ▶ Data from MPI banks (PHOIBLE, “Intercontinental Dictionary Series”), paper dictionaries (and web scraping...)
 - ▶ Tools: phylogenetic software, `lingpy`
 - ▶ Combine known data and phylogeny to infer best models and the order of ancestral state changes
- ▶ Some bold, dreamy research questions:
 - ▶ How confident can we be about the presence and values of laryngeals in PIE? What is the most likely date for their evolution? How many there were? (quantitative criticism of Voyles, 2015)
 - ▶ How confident can we be of Etruscan as a colonial Luwian language? (cf. Woudhuizen, 2008)
 - ▶ How far back can we confidently go in language reconstruction?

