QMSS16 Presentation: Distinct Phonetic Features for Language Phylogeny

Tiago Tresoldi

May 18th, 2016

The questions and the problems

- Is it possible (and useful) to integrate phonology in language phylogeny?
  - Phonology has a high rate of evolution, but borrowing is rare
  - Diachronic and synchronic data allow to estimate the probability of future changes
- Hypothesis: using distinct features ("the most basic unit of analysis in phonetic structures"), adapting Roman Jakobson *et al.* (1941–1956), Chomsky and Halle (1968–1983), and, in particular, Ladefoged (1971; 1993)
  - Pros: binary (and mostly non-exclusive) features (no dummy variables!), single description for vowel 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

The data

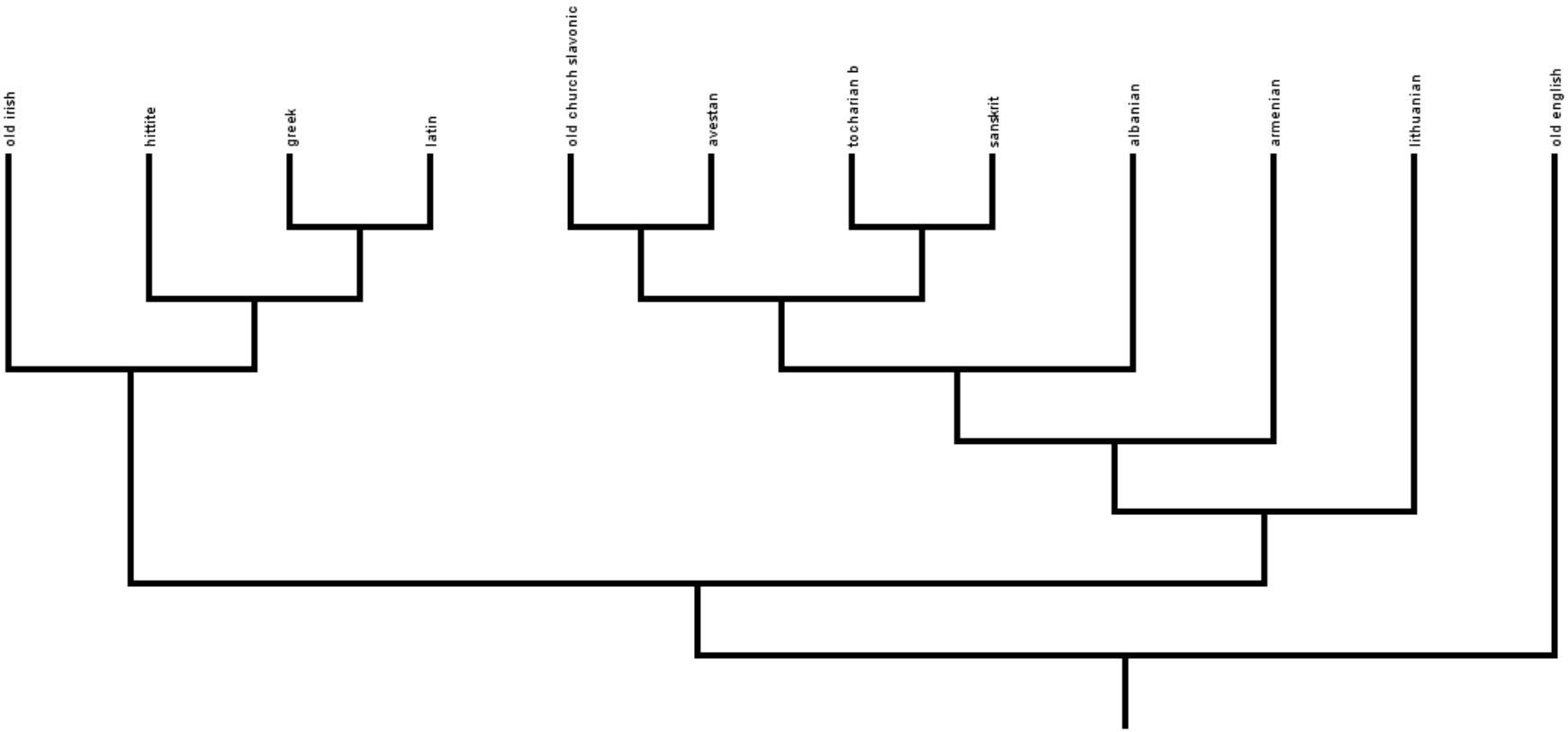
Table: Phonetic reflexes (snippet)

PIE	Old Irish	Latin	Old English	Lithuanian
*p	∅	p	f	p
*ǵʰi	g	h	g	ʒ
*kʷ	k	kʷ	hʷ	k

Table: Distinct features of phonetic reflexes (snippet)

	p.ant	p.obstr	p.lab	ǵʰi.obstr	ǵʰi.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



Conclusion

- Without constrains and additional data, the tree is mirroring similarities in phoneme inventories, not language history
- Ancestral State Recostruction – do results from past methods match those of current ones?
  - Data from MPI banks (PHOIBLE, "Intercontinental Dictionary Series", "World Loanword Database"), and paper dictionaries
  - Tools: phylogenetic software, lingpy
  - Combine known data and phylogeny to infer best model and order of ancestral state changes
- Code on GitHub (`https://github.com/tresoldi/qmss2016`)
- Dreamy questions:
  - How confident can we be on the presence of laryngeals in PIE? What is the most likely date for their evolution? How many there were?
  - How confident can we be of Etruscan as a colonial Luwian language? (cf. Woudhuizen, 2008)