

Extending Parametric Comparison

James Baker, Theoretical and Applied Linguistics (*joint work with Ian Roberts*)

Following work by Longobardi, Guardiano and colleagues on *nominals*, applying methods developed in evolutionary biology:

⇒ **AIM:** to computationally generate **linguistic phylogenies** from **parameters of clausal syntax**

88 parameters in (currently) 33 languages:

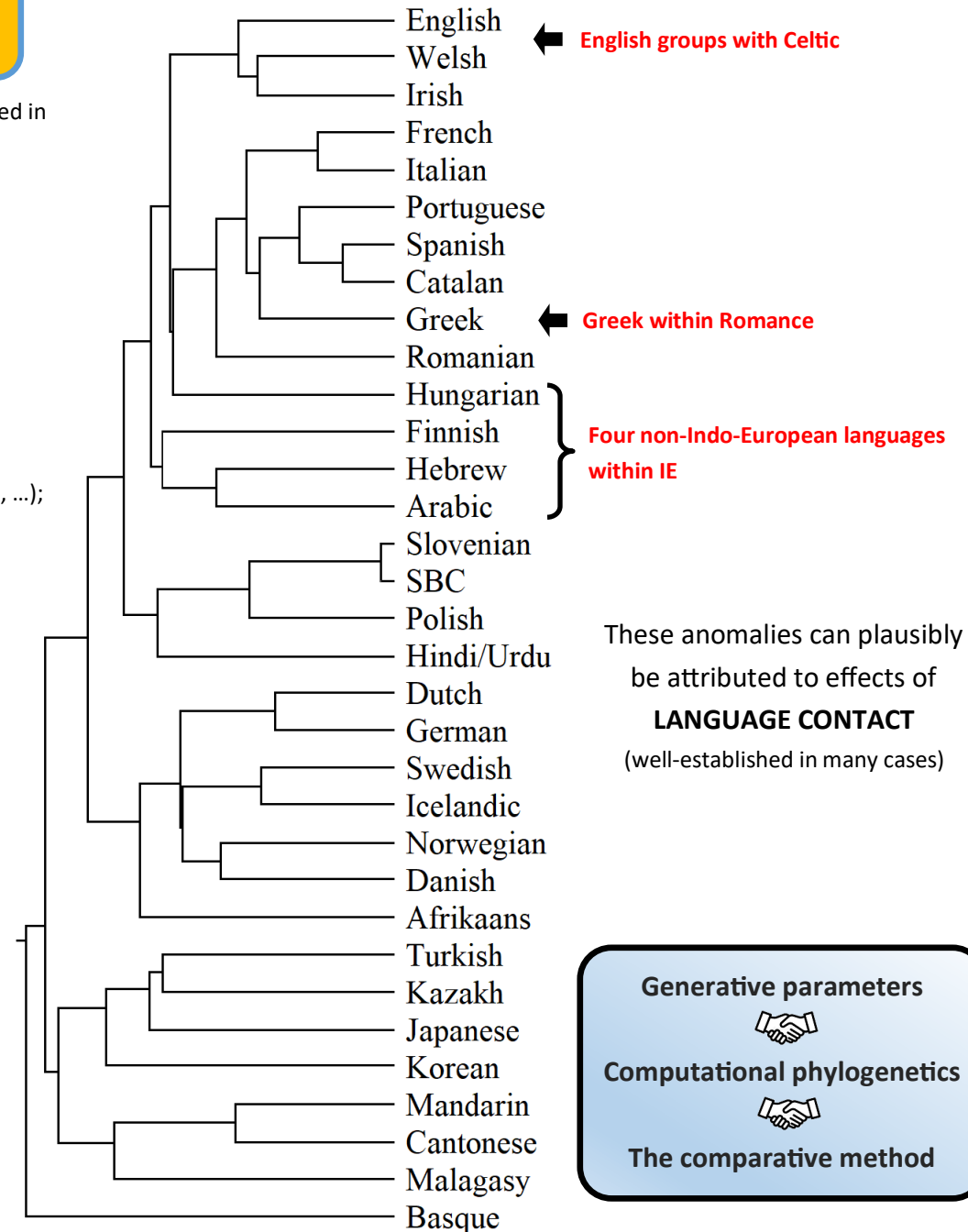
- ◆ Assuming a broadly *generative, minimalist* framework: fine-grained, discrete descriptions of linguistic variation
- ◆ Salient patterns of variation within the clause: e.g.
 - * **presence/absence of grammatical categories** (person, tense, evidentiality, passive voice, ...);
 - * **locus of feature realisation** (e.g. is tense marked on verbs?, ...);
 - * **word order and movement** (e.g. verb-object ordering, *wh*-movement, ...)
- ◆ Each parameter in each language valued either **+**, **-**, or **0**
 - * **0** values predictable from other values, e.g. **+ Grammaticalised Tense** ⇒ **0 Future Tense**

Sample of
parameter
values

#	Code	Descriptor	English	French	Hindi/Urdu	Mandarin
Pv1	VGP	Grammaticalised Person in EP(V)	+	+	+	-
Pv2	VGN	Grammaticalised Number in EP(V)	+	+	+	-
Pv3	VGG	Grammaticalised Gender in EP(V)	-	+	+	-
Pv4	PCV	Φ-feature checking on V	+	+	+	0-
Pv5	PSV	Φ-feature spread to V	-	-	+	0-
Pv6	SAG	Secondary agreement	-	-	-	0-
Pv7	SCL	Subject clitic distinct from agreement	-	-?	-	0-
Pv8	SCE	Subject clitic enclisis	0-	0-	0-	0-
Pv9	GRT	Grammaticalised Tense	+	+	+	-
Pv10	GPT	Grammaticalisation of Past	+	+	+	0-
Pv11	GFT	Grammaticalisation of Future	+	+	+	0-

- ◆ From parameter values we can calculate a **syntactic distance** for each pair of languages (using *Jaccard formula*)
 - * **matrix of all distances** forms the input to phylogenetic programs

Output phylogeny (UPGMA method) **largely corresponds to traditional family trees**, with some **anomalies**:



Generative parameters
Computational phylogenetics
The comparative method

Nb. method requires a **single root node**: we are *not* claiming these languages are all related!