Language history and parameter setting

European Research Council
Established by the European Commission

Andrea Ceolin, Guido Cordoni, Cristina Guardiano, Monica Alexandrina Irimia, Dimitar Kazakov, Shin-Sook Kim, Giuseppe Longobardi, Dimitris Michelioudakis, Nina Radkevich

Language history and parameter setting

1. The Parametric Comparison Method (PCM)
2. Linguistic Theory and the calculations of language distances
3. Phylogenetic Trees
4. The implication for Parameter Theory and Parameter Setting

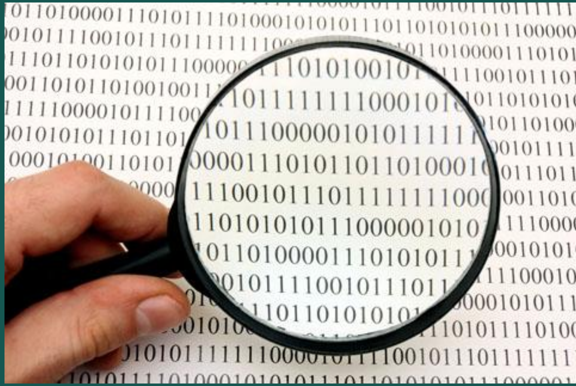
1

The Parametric Comparison Method (PCM)

Guardiano and Longobardi (2005), Longobardi and Guardiano (2009)

1. The Parametric Comparison Method (PCM)

- Parameter values can be used as *comparanda* for historical reconstruction
- Syntactic diversity can be quantified as a distance measure
- Computational taxonomies purely based on syntax can be generated and validated

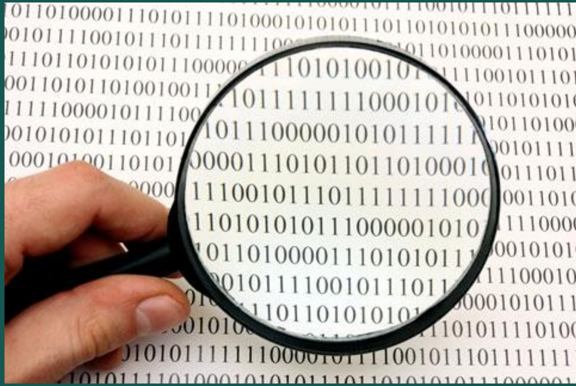


Longobardi (2003), Guardiano and Longobardi (2005), Longobardi and Guardiano (2009)

1. The Parametric Comparison Method (PCM)

The PCM has been successfully used to explore:

- The Historical classification of Indo-European languages
- The Historical classification of 28 languages spoken in Eurasia
- The analysis of syntactic microvariation in Southern Italy



Longobardi et al. (2013), Guardiano et al. (2016), Longobardi et al. *forthcoming*

1. The Parametric Comparison Method (PCM)

A crosslinguistic syntactic difference is a binary parameter if and only if it entails:

- The presence of obligatorily formal expression for a semantic or morphological distinction (the obligatory valuing of an interpretable or uniterpretable feature)
- The variable form of a category depending on a syntactic context (Selection and Feature Agreement)
- The position of a category (Overt Movement)

Guardiano and Longobardi (2016)

1. The Parametric Comparison Method (PCM)

Our data encode properties of the DP, such as:

- ❖ The status of features associated with D, e.g. *person, number, gender and definiteness*
- ❖ Syntactic properties of adjectives, relative clauses, genitival arguments and possessives, demonstratives
- ❖ Type and scope of N-movement

Guardiano and Longobardi (2016)

2

Linguistic Theory and the calculations of language distances

2. Linguistic Theory and the calculations of language distances

A formal model: Principles and Parameters Theory

‘The P&P model is in part a bold speculation rather than a specific hypothesis. Nevertheless, its basic assumptions seem reasonable.... and they do suggest a natural way to resolve the tension between descriptive and explanatory adequacy’

Chomsky (1995)

2. Linguistic Theory and the calculations of language distances

In order to quantify syntactic distances, one can simply count differences in setting:

d(A,B) = 2/3
d(B,C) = 1/3
D(A,C) = 1/3

This is a **Hamming** distance

	A	B	C
P1	+	-	+
P2	+	-	-
P3	-	-	-

Longobardi and Guardiano (2003) and Guardiano and Longobardi (2005)

2. Linguistic Theory and the calculations of language distances

A problem: implications (e.g. grammaticalization of functional projections)

d(A,B) = 1/2, not 2/3
d(B,C) = 1/2, not 2/3
D(A,C) = 1/3

	A	B	C
P1	+	+	+
P2	+	-	+
[+P2] P3	+	0	-

Longobardi and Guardiano (2003) and Guardiano and Longobardi (2005)

2. Linguistic Theory and the calculations of language distances

‘[…] the notion of parametric dependencies runs into empirical problems that should cast doubt on the feasibility of parametric approaches to UG.’

Boeckx and Leivada (2013), Longobardi et al. (2015)

	TABLE A			It	Fr	Ptg	E	D	SC	Po	Rus	Ma	Hi	Ar	Heb	Hu	Fin	Wo	
1	± gramm. person		FGP	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	1
2	± gramm. number	+FGP	FGN	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	2
3	± gramm. gender	+FGN	FGG	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	3
4	± NP over D	+FGP	NOD	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	4
5	± feature spread to N	+FGN	FSN	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	5
6	± numb. on N (BNs)	+FSN	FNN	+	-	+	+	+	+	+	+	+	+	+	+	+	+	0	6
7	± gramm. partial def		DGP	+	+	+	+	+	-	-	-	-	-	+	+	+	-	+	7
8	± gramm. def	+DGP	DGR	+	+	+	+	+	0	0	0	0	0	+	+	+	0	+	8
9	± strong person	+FGP, +DGR, -NOD	NSD	+	+	+	-	-	0	0	0	0	0	+	+	+	0	0	9
10	± free null partitive Q	+FNN	DPQ	-	0	-	-	-	-	-	-	-	-	-	-	-	+	0	10
11	± gramm. dist. art.	-FSN or -FNN or +DGR	DDA	-	-	-	-	-	0	0	0	0	0	-	-	-	0	+	11
12	± def-checking N	+DGR	DCN	-	-	-	-	-	0	0	0	0	0	-	-	-	0	-	12
13	± def spread to N	+DCN, -NSD	DSN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13
14	± def on relatives	+DGR	DOR	-	-	-	-	-	0	0	0	0	0	+	-	-	0	+	14
15	± D-controlled infl. on N	+FSN	DIN	-	-	-	-	-	-	-	-	-	-	+	-	-	-	0	15
16	± plural spread from cardinals	+FSN	CPS	+	+	+	+	+	+	+	+	+	+	+	+	-	-	0	16
17	± gramm. boundedness		CGB	-	-	-	-	-	-	-	-	+	+	-	-	+	-	+	17
18	± strong article	+DGR, +FNN, -CGB	CGR	+	0	+	+	+	0	0	0	0	0	-	-	0	0	0	18
19	± bounded.-checking N	+CGB	CCN	0	0	0	0	0	0	0	0	-	-	0	0	-	0	-	19
20	± null-N-licensing art	-FSN or-FNN or -DCN, +NOD or +NSD	DNN	-	-	+	0	0	0	0	0	0	0	-	-	-	0	?	20
21	± structured APs		AST	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	21
22	± feature spread to struct. APs	+FSN, + AST	FFS	+	+	+	-	+	+	+	+	+	+	+	+	-	+	0	22
23	± feature spread to pred. APs	+FGN	FSP	+	+	+	-	-	+	+	+	+	+	+	+	+	+	+	23
24	± D-controlled infl. on A	-NSD, +FFS	ADI	0	0	0	0	+	0	0	0	0	0	0	0	0	0	0	24
25	± DP over relatives		ADR	+	+	+	+	+	+	+	+	-	-	+	+	+	+	+	25
26	± relative extrap.	-ADR	AER	0	0	0	0	0	0	0	0	+	+	0	0	0	0	0	26
27	± free reduced rel	+AST	ARR	+	+	+	-	-	-	-	-	+	+	+	-	-	-	0	27
28	± N-raising with obl. pied-piping	+AST	NPP	-	-	-	-	-	-	-	-	-	-	+	+	-	-	0	28
29	± free Gen		GFR	+	+	+	+	+	-	-	-	+	+	+	+	-	-	-	29
30	± uniform Gen	+GFR	GUN	-	-	-	-	-	0	0	0	-	-	-	-	0	0	0	30
31	± DP over free Gen	+GFR, +ADR	GPR	+	+	+	+	+	0	0	0	0	0	+	+	0	0	0	31
32	± GenO	≈+GUN	GFO	-	-	-	-	+	+	+	+	-	-	+	+	-	-	+	32
33	± Gen-feature spread to N		GFS	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	33
34	± D checking poss.	+DGR, +NSD or ≈+CGR	PDC	-	+	?	0	0	0	0	0	0	0	-	-	-	0	-	34
35	± adjectival poss.		APO	+	+	+	-	+	+	+	+	-	-	-	-	-	-	-	35
36	± post-affix poss.	+DCN	PAP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	36
37	± clitic poss.		PCL	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	37
38	± N-feat. spr. to pron. poss.	+FFS or -AST, +PAP or +PCL	PHS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-	38
39	± N-feature spread to free Gen	+FFS, +GFR, ≈-PHS	GSP	-	-	-	0	-	0	0	0	+	+	-	-	0	0	0	39
40	± adjectival Gen	+APO	AGE	-	-	-	0	-	+	-	+	0	0	0	0	0	0	0	40
41	± Poss°-checking N	-GFS	GCN	-	-	-	-	-	-	-	-	-	-	+	+	0	0	?	41
42	± Strong partial Locality	≈+CGR, -FSN or +FNN	TPL	0	0	0	0	0	+	+	+	+	+	+	-	+	+	+	42
43	± Strong Locality	≈-TPL	TSL	+	+	+	+	+	+	+	+	+	+	-	0	+	+	+	43
44	± D Checking Dem	≈-TPL, -FSN or +DGR	TDC	+	+	+	+	+	0	0	0	0	0	-	0	-	0	+	44
45	± N over cardinals		NOC	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	45
46	± N over ordinals	-NOC	NOO	-	-	-	-	-	-	-	-	-	-	0	0	-	-	-	46
47	± N over M1 As	-NOO, -NPP	NM1	-	-	-	-	-	-	-	-	-	-	0	0	-	-	0	47
48	± N over M2 As	-NM1	NM2	+	+	+	-	-	-	-	-	-	-	0	0	-	-	0	48
49	± N over As	-NM2	NOA	0	0	0	-	-	-	-	-	-	-	0	0	-	-	0	49
50	± N over GenO	≈-GFO, -NOA or -AST	NGO	0	0	0	0	+	+	+	+	0	0	0	0	0	0	+	50
51	± N over ext. arg.	-NGO or (-GFO, -NOA or -AST)	NOE	0	0	0	-	0	0	0	0	+	+	0	0	+	+	0	51
52	± free MOD	-NOA	AFM	0	0	0	-	-	-	-	-	-	-	0	0	-	-	0	52
53	± class MOD	-AFM	ACM	0	0	0	-	-	-	+	+	-	-	0	0	-	-	0	53
54	± def on APs	+DGP, *+postnom. APs*	DOA	-	-	-	-	0	0	0	0	0	0	+	+	0	0	-	54
55	± gramm. AP marker	*+postnom. APs*	AMO	-	-	-	-	0	0	-	-	0	0	-	-	0	0	+	55
56	± Cons. Pr.	(-NM1,*+A-Cpl*) or (+NPP or ≈-NM2, *+Cpl-A*)	ACP	+	+	+	+	+	-	-	-	0	0	0	0	+	+	0	56
				It	Fr	Ptg	E	D	SC	Po	Rus	Ma	Hi	Ar	Heb	Hu	Fin	Wo	

2. Linguistic Theory and the calculations of language distances

The Borer-Chomsky conjecture

"All parameters of variation are attributable to differences in the features of particular items (e.g., the functional heads) in the lexicon."

Borer (1984), Chomsky (1995), Baker (2008)

2. Linguistic Theory and the calculations of language distances

Parameter Schemata

- Is F, F a feature, **grammaticalized**?
- Does F, F a grammaticalized feature, **Agree** with X, X a category (i.e. does F probe X)?
- Is F, F a grammaticalized feature, **spread** on X, X a category?
- Is F, F a grammaticalized feature, “strong” (i.e. does F **overtly attract** X, probe X with an EPP feature)?
- Does a functional category (a set of lexically co-occurring grammaticalized features) X have a **phonological matrix Φ**?
- Is F, F a grammaticalized feature, **checked** by the minimal accessible category of type X (or is pied-piping possible)?

Longobardi (2005)

2. Linguistic Theory and the calculations of language distances

In order to quantify syntactic distances, one should only consider the number of lexical features:

d(A,B) = 2/2
d(B,C) = 1/1
D(A,C) = 1/2

This is a **Jaccard** distance (identities in '–' are not counted)

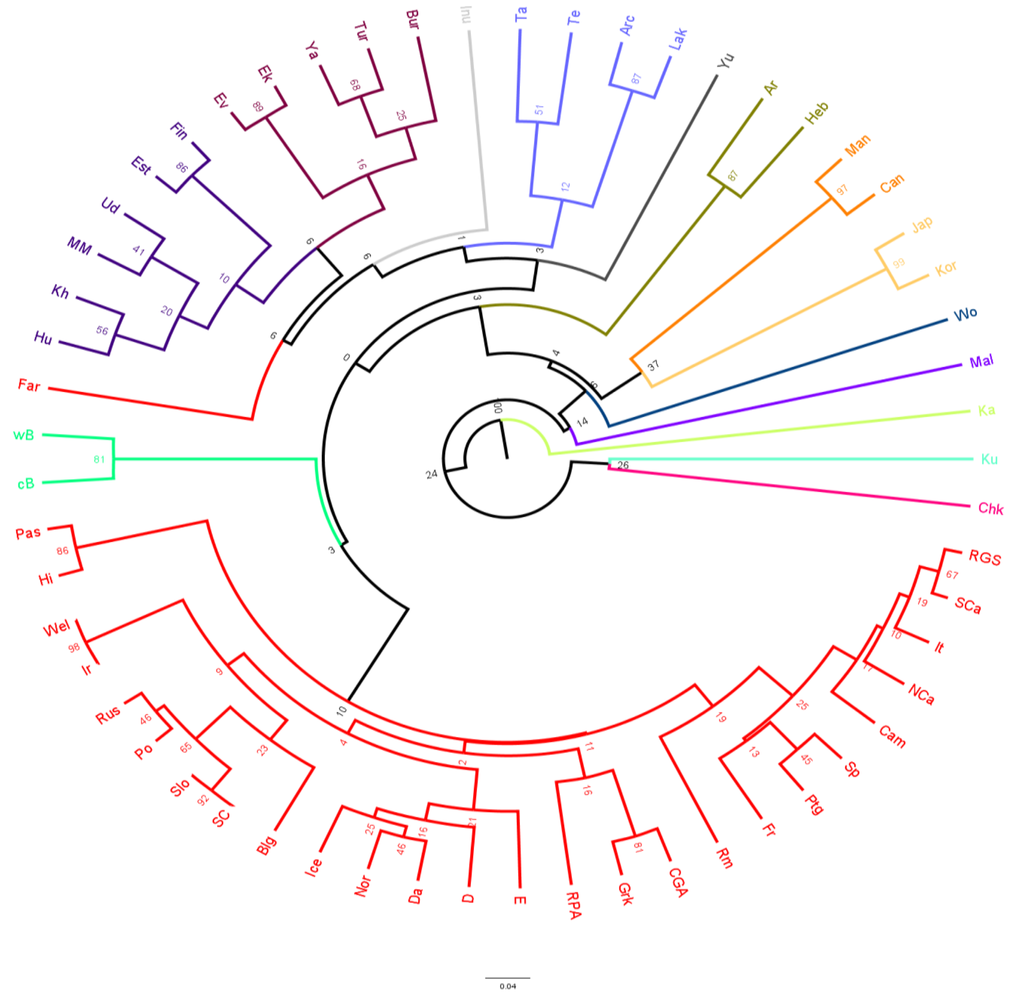
	A	B	C
P1	+	-	+
P2	+	-	-
P3	-	-	-

3

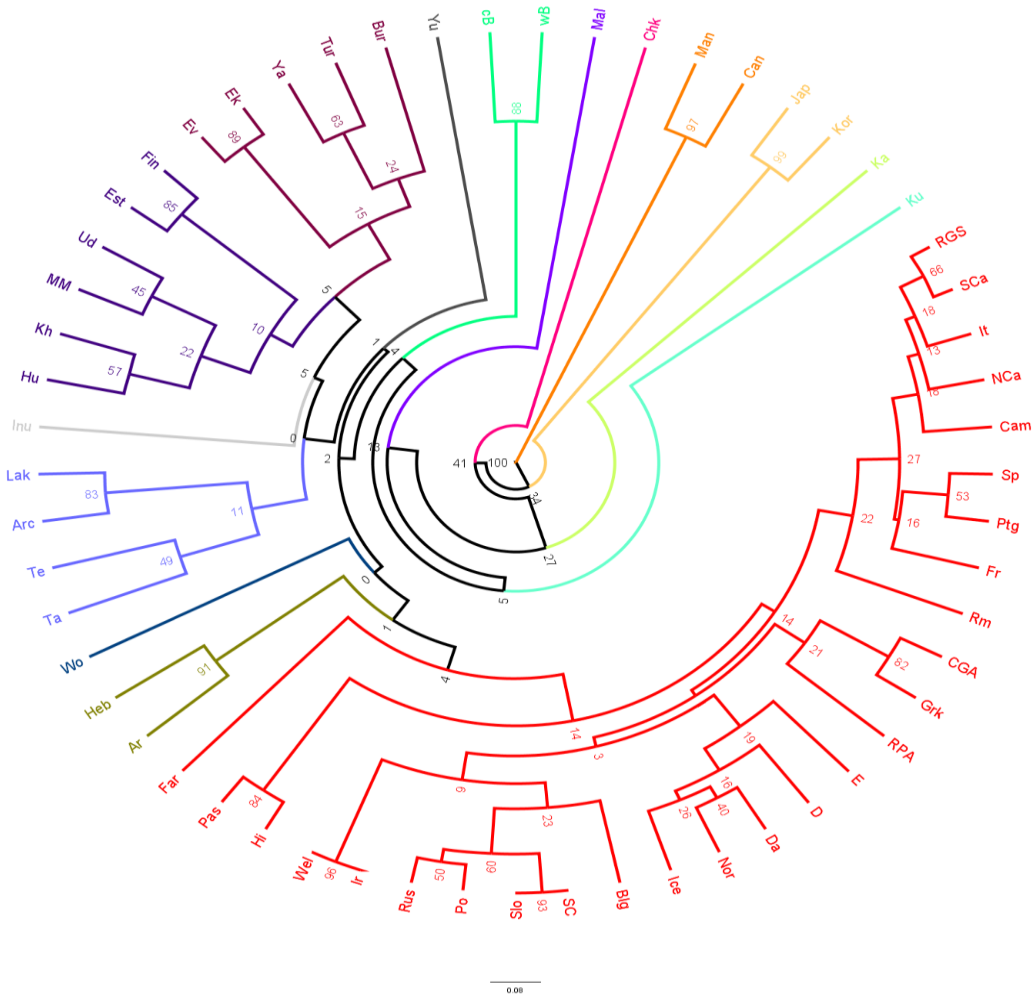
Phylogenetic trees

Longobardi and Guardiano (2009), Longobardi et al. (2013)

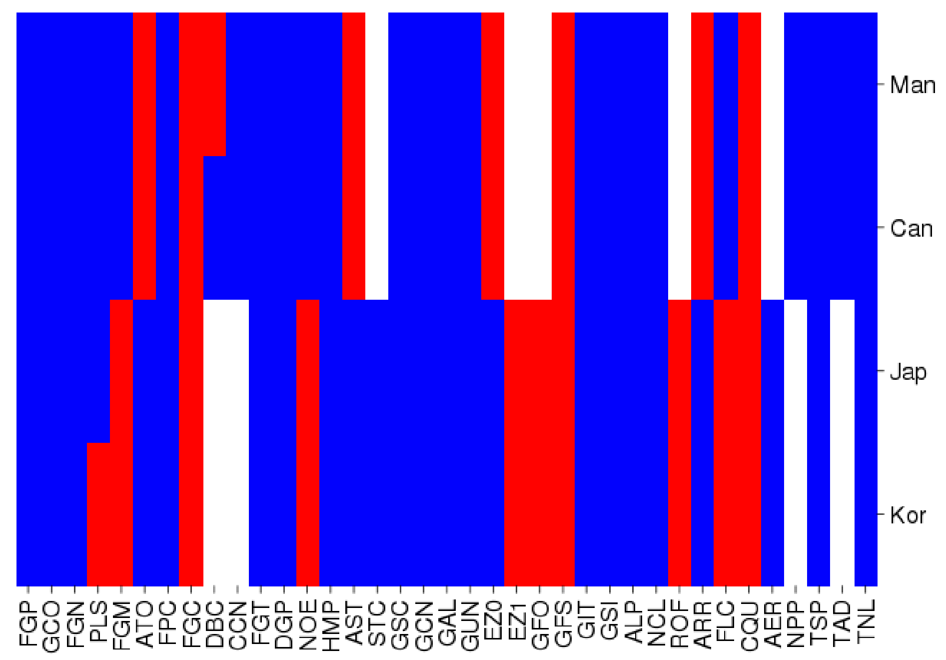
3. Phylogenetic trees



3. Phylogenetic trees



3. Phylogenetic trees



3. Phylogenetic trees

	Jap	Kor	Man	Can
FGP, gr. person	-	-	-	-
FGN, gr. number	-	-	-	-
NCL, clitic poss.	-	-	-	-

3. Phylogenetic trees

	Hi	Far	Tur	Ar
SGE, sem. gender	?	-	-	?
ROF, Rel over N	?	-	+	?
PSC, Card. F. spread	?	-	-	?

4

The implication for Parameter Theory and Parameter Setting

4. The implication for Parameter Theory and Parameter Setting

- A first step towards simplification of parameter setting was the formulation of an 'implicational structure', for which parameters whose value can be deduced ('0') were removed from the computation, since they are not part of the mental representation of the language.

- Here we make a further step by claiming that also parameters which are set on ('–') are not always relevant, because they represent the 'absence' of a feature. They are only relevant when we count differences in the number of features, not identities.

- The only thing which is truly shared by languages are '+' values, which represent the presence of a feature.

4. The implication for Parameter Theory and Parameter Setting

❖ The fact that this system provides more plausible phylogenetic results is a proof of the representation of grammars as lists of syntactic features.

THANKS!

Any questions?

ceolin@sas.upenn.edu
monica_alexandrina@yahoo.com