Typologies in equilibrium

Gerhard Jäger

Tübingen University

University of Zurich, March 15, 2018

WORDS BONES GENES TOOLS
Tracking Linguistic, Cultural, and Biological Trajectories of the Human Past

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

DFG

erc
European Research Council
Established by the European Commission

Case alignment systems

Universal syntactic-semantic primitives

- three universal core roles
  - **S:** intransitive subject
  - **A:** transitive subject
  - **O:** transitive object

German

Der Junge	ist	dreckig.
the boy.NOM	is	dirty

'The boy is dirty.'

Der Junge	wirft	einen Stein.
DEF boy.NOM	throw	a.ACC stone

'The boy is throwing a stone.'

S
A
O

Kalkatungu (Australia)

Kaun	muu-yan-ati
dress.ABS	dirt-PROP-INCH

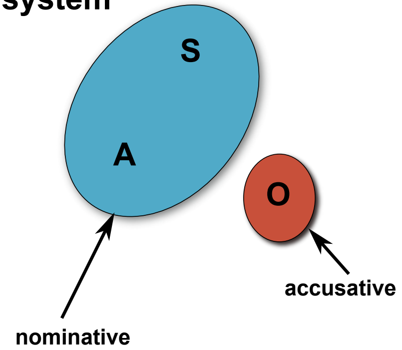
'The dress is dirty.'

Kuntu	wampa-ngku	kaun	muu-yan-puni-mi.
not	girl-ERG	dress.ABS	dirty-PROP-CAUS-FUT

'The girl will not dirty the dress.'

Alignment systems

Accusative system



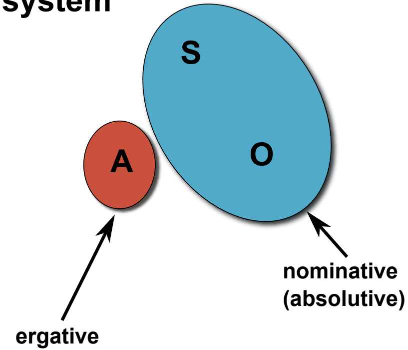
Latin

Puer puellam vidit.
boy.NOM girl.ACC saw *'The boy saw the girl.'*

Puer venit.
boy.NOM came *'The boy came.'*

Alignment systems

Ergative system



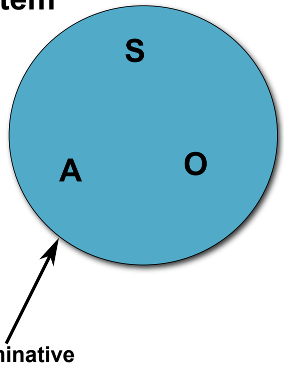
Dyirbal

ŋuma yabu-ŋgu bura-n.
father mother.ERG see-NONFUT
'The mother saw the father.'

ŋuma banaga-nu.
boy.NOM came *'The boy came.'*

Alignment systems

Neutral system



Mandarin

rén lái le.
person come CRS
'The person has come.'

zhāngsān mà lǐsì le ma.
Zhangsan scold Lisi CRS Q
'Did Zhangsan scold Lisi?'

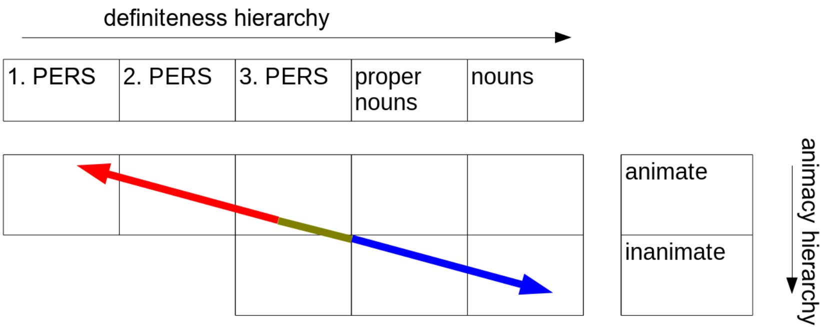
Differential case marking

- many languages have mixed systems
- e.g., some NPs have accusative and some have neutral paradigm, such as Hebrew

(1) Ha-seret her?a ?et-ha-milxama
the-movie showed acc-the-war
'The movie showed the war.'

(2) Ha-seret her?a (*?et-)milxama
the-movie showed (*acc-)war
'The movie showed a war'
(from Aissen, 2003)

Differential case marking



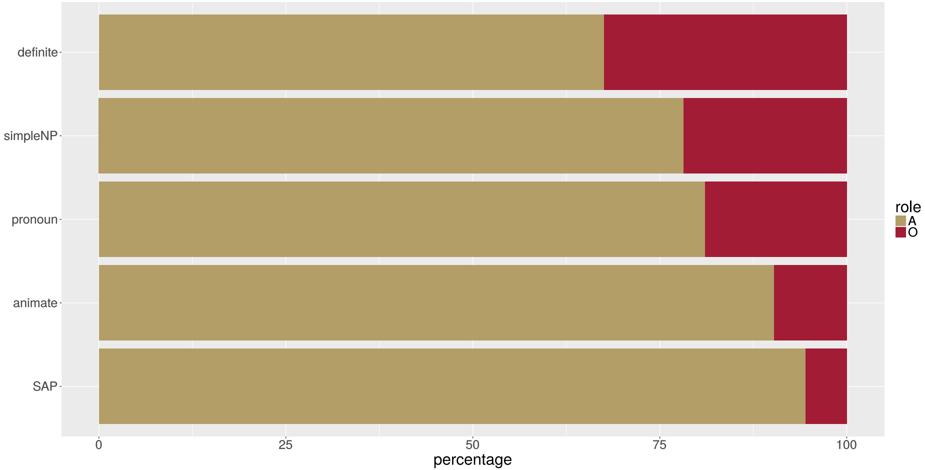
accusative

neutral or tripartite

ergative

Functional explanation?

probability P(syntactic role|prominence of NP)



A note on terminology

A is prominent	A is non-prominent	O is prominent	O is non-prominent
e(rgative)	e(rgative)	a(ccusative)	a(ccusative)
e	e	a	z(ero)
e	e	z	a
e	e	z	z
e	z	a	a
...	...	...	...
z	e	z	z
z	z	a	a
z	z	a	z
z	z	z	a
z	z	z	z

A note on terminology

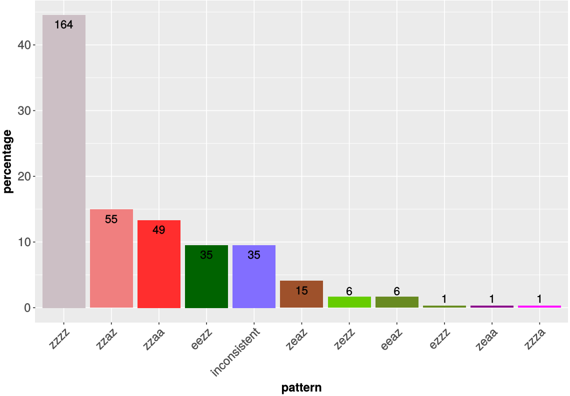
actually attested:

1. **zzzz**: no case marking
2. **zzaa**: non-differential object marking
3. **zzaz**: harmonic differential object marking
4. **ezzz**: non-differential subject marking
5. **zeaz**: split ergative
6. **eeaz**: non-differential subject marking plus differential object marking
7. **ezzz**: dis-harmonic differential subject marking
8. **zezz**: harmonic differential subject marking
9. **zeaa**: harmonic differential subject marking plus non-differential object marking
10. **zzza**: dis-harmonic differential object marking

Empirical distribution

Bickel et al.'s (2015) sample

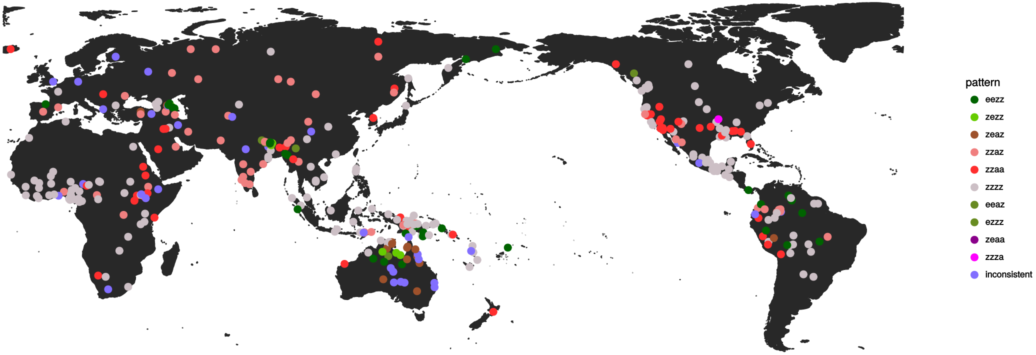
- genetically diverse sample of 460 case marking systems
- used here: 368 systems
  - one system per language
  - only languages with ISO code
  - only languages present in ASJP
- 2 out of 333 systems (99.4%) are obey the Silverstein hierarchy (not counting inconsistent states)



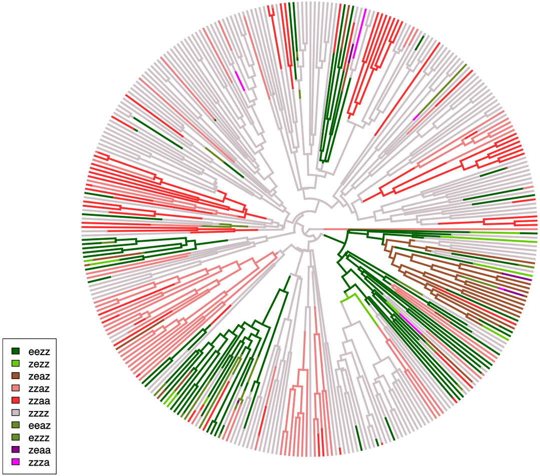
Phylogenetic non-independence

- languages are phylogenetically structured
- if two closely related languages display the same pattern, these are not two independent data points

⇒ we need to control for phylogenetic dependencies



Phylogenetic non-independence



Phylogenetic non-independence

Maslova (2000):

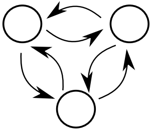
"If the A-distribution for a given typology cannot be assumed to be stationary, a distributional universal cannot be discovered on the basis of purely synchronic statistical data."

*"In this case, the only way to discover a distributional universal is to **estimate transition probabilities** and as it were to 'predict' the stationary distribution on the basis of the equations in (1)."*

The phylogenetic comparative method

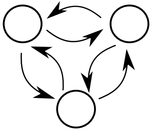
Modeling language change

Markov process

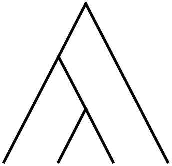


Modeling language change

Markov process

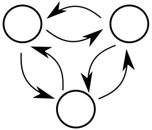


Phylogeny

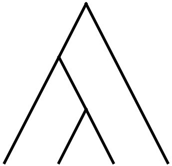


Modeling language change

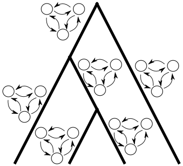
Markov process



Phylogeny

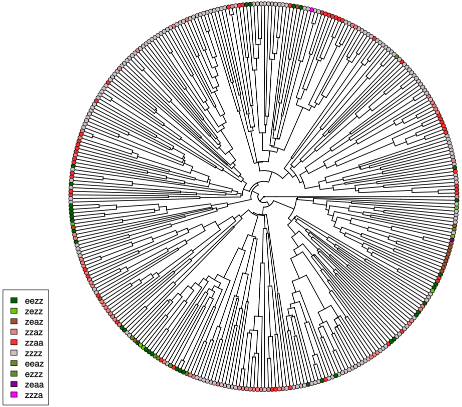


Branching process



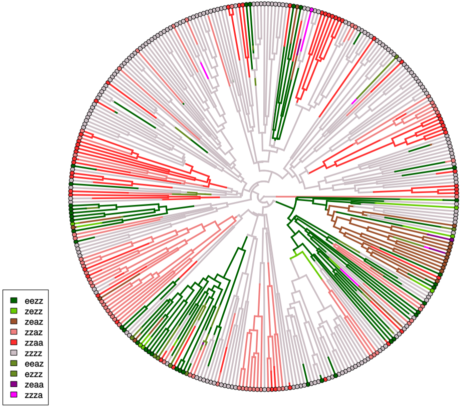
Estimating rates of change

- if phylogeny and states of extant languages are known...



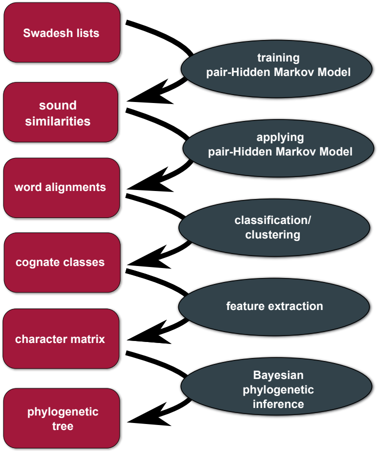
Estimating rates of change

- if phylogeny and states of extant languages are known...
- ... transition rates and ancestral states can be estimated based on Markov model

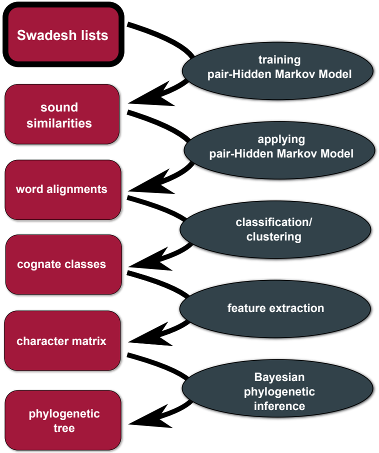


Inferring a world tree of languages

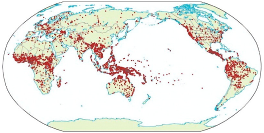
From words to trees



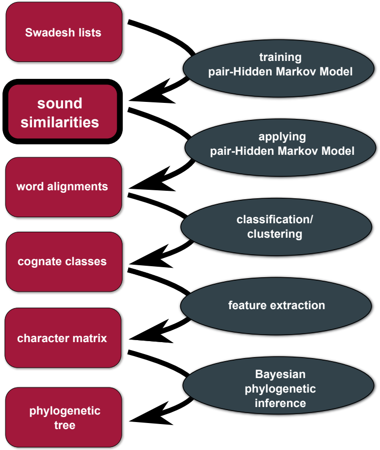
From words to trees

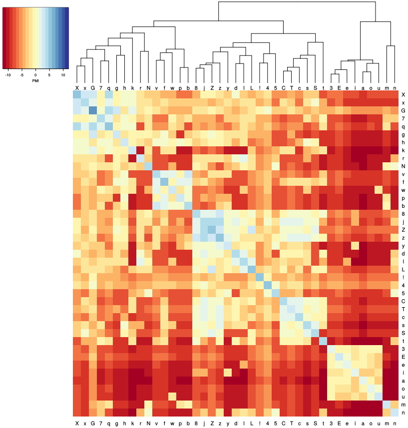


concept	Latin	English
I	ego	Ei
you	tu	yu
we	nos	wi
one	unus	w3n
two	duo	tu
person	persona, homo	pers3n
fish	piskis	fiS
dog	kanis	dag
louse	pedikulus	laus
tree	arbor	tri
leaf	foly∼u*	lif
skin	kutis	skin
blood	saNgw∼is	bl3d
bone	os	bon
horn	kornu	horn
ear	auris	ir
eye	okulus	Ei

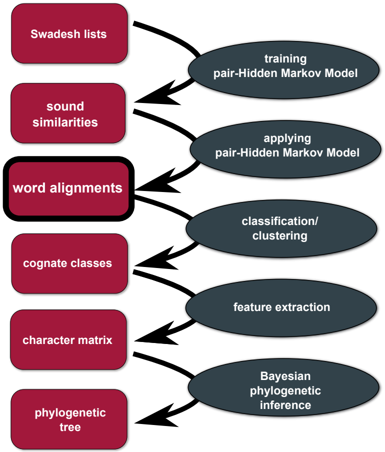


From words to trees



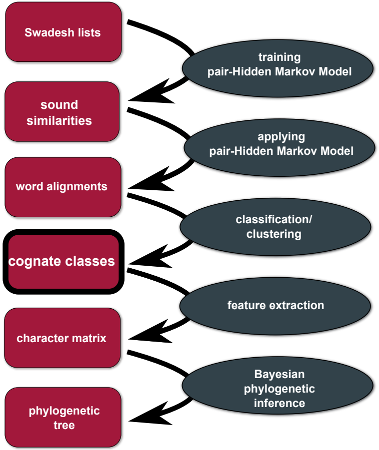


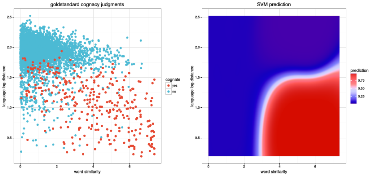
From words to trees



Language	*fish:z*	*tongue:1*	*smoke:1*
Abui-Atangmelang	-af-u		
Abui-Fuimelang	-af-u	tal-i-fi--	
Adang	aab--	tal-E-b---	awai--b-a-n-o-7o-
Blagar-Bakalang	-ab--	--j-e-bur-	--ad--b-a-n-aNka-
Blagar-Bama	aab--	teg-e-bur-	------b-e-n-a-xa-
Blagar-Kulijahi	-ab--	tej-e-bur-	------b-e-n-aNka-
Blagar-Nule	aab--	tej-e-bur-	--ad--b-e-n-aNka-
Blagar-Tuntuli	aab--	tej-e-bur-	a-adgeb-a-n-a-q--
Blagar-Warsalelang	-ab--	tel-e-bur-	a-ad--b-a-n-a-x--
Bunaq			------b-o-t-o-h--
Deing	haf--		------buu-n------
Hamap	7ab--	nar-ø-buN-	------b-a-n-o-7--
Kabola	hab--	tal-e-b---	awal--b-e-n-e-7o-
Kaera-Padangsul	-ab--	talee-b---	a-ad--b-e-naa-x--
Kafoa	-afUi	tal-i-p---	------f-o-n-a----
Kamang	-ap-i	nal---pu--	------p-u-n----a-
Kiraman	-Eb--	nal-i-bar-	--ar--b-a-n-o-kan
Klon	-eb-i	gel-E-b---	--ed-ab-o-n------
Kui	-eb--	tal-i-ber-	--ar--b-o-n-o-k--
Kula	-ap-i	-il-I-p---	------p---n-ekka-
Nedebang	aaf-i	gel-e-fu--	--ar-ab-u-n------
Reta	aab--	nal-e-bul-	a-ad--b-o-n-a----
Sar-Adiabang	haf--	--p-e-fal-	--ar--buu-n------
Sar-Nule	haf--	nal-e-faj-	
Sawila	-ap-i	gal-impuru	------p-u-n-a-ka-
Teiwa-Madar	xaf--	gel-i-vi--	------buu-n------
Wersing	-ap-i	nej-e-bur-	--ad-ap-u-n-a-k--
Wpantar	hap--	nal-e-bu--	------b-unn-a----

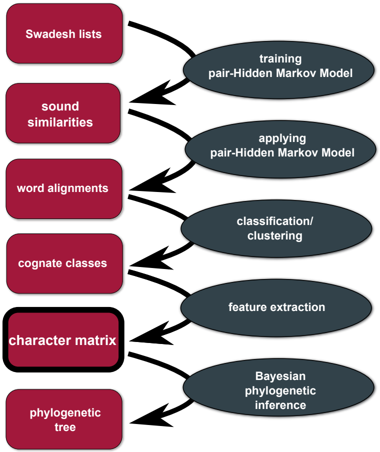
From words to trees





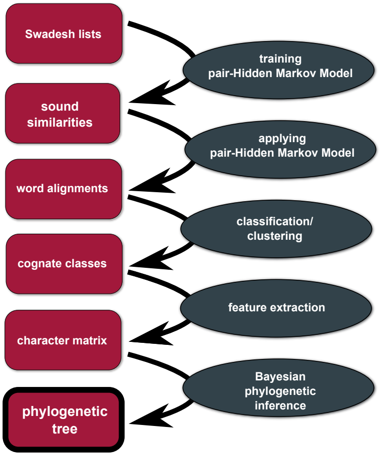
	English	Spanish	Modern Greek	Standard German
I	Ei:A	yo:B	exo:C	iX:D
you	yu:A	ustet:B, tu:C	esi:D	du:E
we	wi:A	nosotros:B	emis:C	vir:A
one	w3n:A	uno:B	enas:C, ena:C	ains:D
two	tu:A	dos:B	8y∼o:C, 8io:D	cvai:E
person	pers3n:A	persona:A	an8∼ropos:B	mEnS:C
fish	fiS:A	peskado:A, pes:A	psari:B	fiS:A
dog	dag:A	pero:B	sTili:C, sTilos:C	hunt:D
come	k3m:A	veni:B	erx∼o:C	kh∼om3n:A
sun	s3n:A	sol:B	ily∼os:C, iLos:C	zon3:A
star	star:A	estreya:A	asteri:A, astro:A	StErn:A
water	wat3r:A	agw∼a:B	nero:C	vas3r:A
stone	ston:A	piedra:B	petra:B	Stain:A
fire	fEir:A	fuego:B	foty∼a:C	foia:D
path	pE8:A	senda:B	8romos:C	pf∼at:A, vek:D
mountain	maunt3n:A	sero:B, monta5a:A	vuno:C, oros:D	bErk:E
full	ful:A	yeno:B	yematos:C, pliris:D	fol:A
new	nu:A	nuevo:A	neos:A, Tenury∼os:B	noi:A
name	nem:A	nombre:A	onoma:A	nam3:A

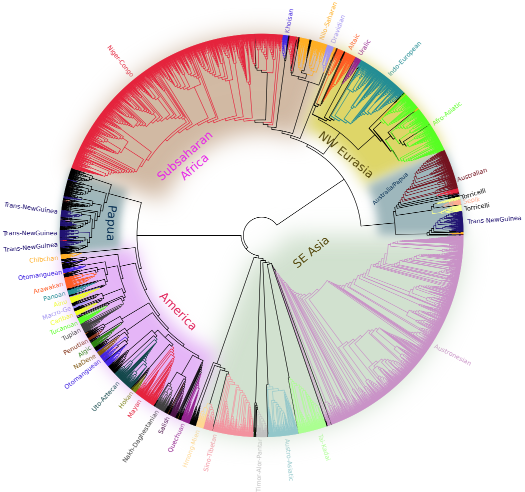
From words to trees




```
TNG.ENGAN.MAIBI                    10000000000000000000000000000000000000000000
TNG.ENGAN.POLE                     00000000000000000000000000000000000010000000
TNG.ENGAN.SAU                      00000000000000000000000000000000000000000000
TNG.ENGAN.YARIBA                   10000000000000000000000000000000000000000000
TNG.FASU.FASU                      00000000000000000000000000000000000010000000
TNG.FASU.NAMUMI                    00000000000000000000000000000000000000001000
TNG.FINISTERRE-HUON.AWARA          00000000000000000000000000000000000010000000
TNG.FINISTERRE-HUON.BORONG         00000000000000000000000000000000000010000000
TNG.FINISTERRE-HUON.BURUM          00000000000000000000000000000000000010000000
TNG.FINISTERRE-HUON.BURUM_MIND     00000000000000000000000000000000001010000000
TNG.FINISTERRE-HUON.DEDUA          00000000000000000000000000000000000010000000
TNG.FINISTERRE-HUON.HUBE           00000000000000000000000000000000000010000000
TNG.FINISTERRE-HUON.KATE           00000000000000000000000000000000000010000000
TNG.FINISTERRE-HUON.KOMBA          00000000000000000000000000000000000010000000
TNG.FINISTERRE-HUON.KOSORONG       00000000000000000000000000000000000010000000
TNG.FINISTERRE-HUON.MAPE           00000000000000000000000000000000000100000000
TNG.FINISTERRE-HUON.MAPE_2         00000000000000000000000000000000000010000000
TNG.FINISTERRE-HUON.MIGABAC        00000000000000000000000000000000000010000000
TNG.FINISTERRE-HUON.MINDIK         00000000000000000000000000000000000010000000
TNG.FINISTERRE-HUON.MOMOLILI       00000000000000000000000000000000000010000000
TNG.FINISTERRE-HUON.NABAK          00000000000000000000000000000000000100000000
TNG.FINISTERRE-HUON.NANKINA        00000000000000000000000000000000000010000000
TNG.FINISTERRE-HUON.NEK            00000000000000000000000000000000000001000000
TNG.FINISTERRE-HUON.NUKNA          00000000000000000000000000000000000001000000
TNG.FINISTERRE-HUON.ONO            00000000000000000000000000000000000010000000
TNG.FINISTERRE-HUON.SELEPET        00000000000000000000000000000000000010000000
TNG.FINISTERRE-HUON.TIMBE          00000000000000000000000000000000000010000000
TNG.FINISTERRE-HUON.TOBO           00000000000000000000000000000000000010000000
TNG.FINISTERRE-HUON.WANTOAT        00000000000000000000000000000000000110000000
TNG.FINISTERRE-HUON.YOPNO          00000000000000000000000000000000000001000000
TNG.GOILALAN.AFOA                  00000000000000000000000000000000000110000000
TNG.GOILALAN.KUNIMAIPA             00000000000000000000000000000000000010000000
TNG.GOILALAN.MAFULU                00000000000000000000000000000000000010000000
```

From words to trees



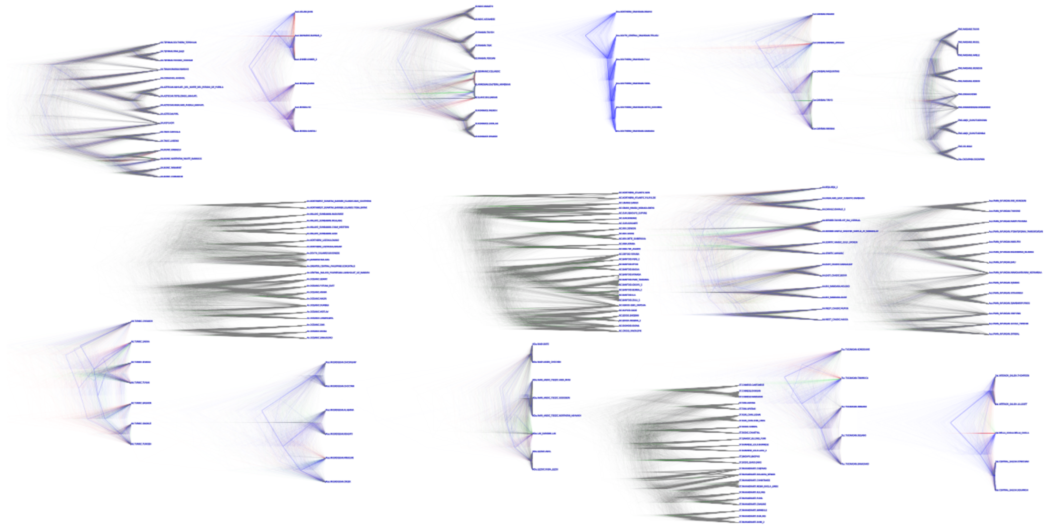


Cases in equilibrium

Phylogenetic trees for the case data

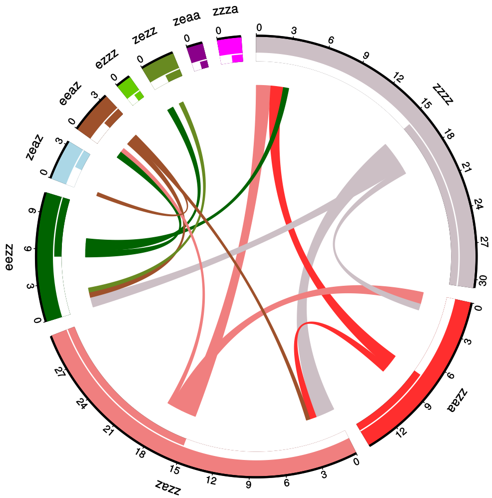
- for 16 Glottolog language families, there are at least five languages in the intersection of Bickel's with the ASJP data (Afro-Asiatic, Atlantic-Congo, Austroasiatic, Austronesian, Cariban, Dravidian, Indo-European, Muskogean, Nakh-Daghestanian, Nuclear Trans-New Guinea, Pama-Nyungan, Salishan, Sino-Tibetan, Tucanoan, Turkic, Uto-Aztecan)
- for each of these families, I inferred a posterior distribution of 1,000 trees (using ASJP character data with MrBayes) to reflect uncertainty in tree structure and branch length
- Glottolog tree was used as constraint tree

Phylogenetic trees for the case data



Estimating transition rates

- totally unrestricted model, all 90 transition rates are estimed independently
- implementation using RevBayes (Höhna et al., 2016)



Reconstructing change events

- using the inferred parameters, it can be estimated how many state changes must have occurred, given what we know about the phylogenies and the states of extant languages
- implemented using SIMMAP algorithm (*stochastic character mapping*, Bollback 2006)

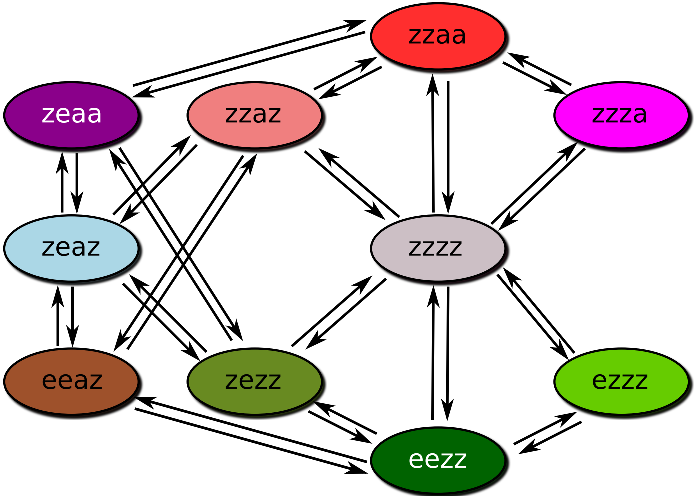
	zzzz	zzaa	zzaz	eezz	zeaz	eeaz	ezzz	zezz	zeaa	zzza
zzzz	-	9.1	25.0	11.6	0.7	1.2	0.4	0.7	0.5	0.5
zzaa	6.9	-	4.1	2.2	0.7	1.6	0.4	0.6	0.4	0.3
zzaz	14.9	5.7	-	3.3	0.6	1.2	0.6	0.6	0.5	0.4
eezz	4.7	2.9	2.8	-	5.0	5.2	0.6	5.9	0.6	0.4
zeaz	1.0	1.3	0.8	4.9	-	1.4	0.3	2.3	2.1	0.3
eeaz	2.4	2.4	2.3	6.6	1.7	-	0.3	0.8	0.3	0.3
ezzz	0.8	0.6	0.7	0.8	0.3	0.4	-	0.4	0.2	0.1
zezz	1.1	0.8	0.8	1.9	0.9	0.8	0.2	-	0.3	0.2
zeaa	0.7	0.7	0.4	0.7	0.9	0.4	0.2	0.4	-	0.1
zzza	0.4	0.5	0.6	0.6	0.3	0.4	0.2	0.4	0.2	-

Refining the model

These numbers suggest that almost all inferred changes are of the following type:

- gaining or losing object marking
- gaining or losing subject marking
- gaining or losing object marking on prominent objects
- gaining or losing subject marking on prominent objects
- gaining or losing object marking on non-prominent objects
- gaining or losing subject marking on non-prominent objects

Refining the model

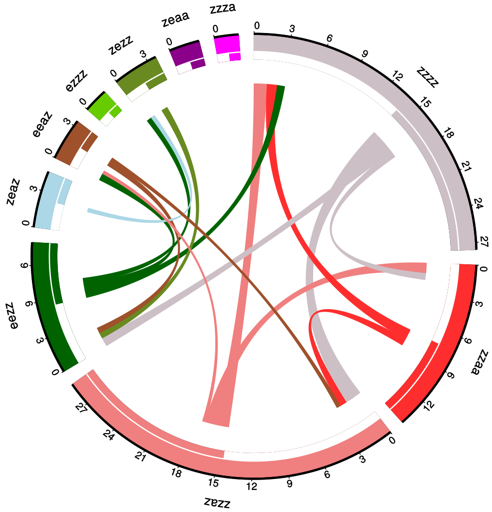


Refining the model

	zzzz	zzaa	zzaz	eezz	zeaz	eeaz	ezzz	zezz	zeaa	zzza
zzzz	-	9.1	25.0	11.6	0.7	1.2	0.4	0.7	0.5	0.5
zzaa	6.9	-	4.1	2.2	0.7	1.6	0.4	0.6	0.4	0.3
zzaz	14.9	5.7	-	3.3	0.6	1.2	0.6	0.6	0.5	0.4
eezz	4.7	2.9	2.8	-	5.0	5.2	0.6	5.9	0.6	0.4
zeaz	1.0	1.3	0.8	4.9	-	1.4	0.3	2.3	2.1	0.3
eeaz	2.4	2.4	2.3	6.6	1.7	-	0.3	0.8	0.3	0.3
ezzz	0.8	0.6	0.7	0.8	0.3	0.4	-	0.4	0.2	0.1
zezz	1.1	0.8	0.8	1.9	0.9	0.8	0.2	-	0.3	0.2
zeaa	0.7	0.7	0.4	0.7	0.9	0.4	0.2	0.4	-	0.1
zzza	0.4	0.5	0.6	0.6	0.3	0.4	0.2	0.4	0.2	-

Refining the model

- next step: re-estimate parameters with *a priori* assumptions that only the 36 (out of 90) rates consistent with the generalizations are > 0
- Bayesian model comparison: ΔAICM $= 20.8$ in favor of the restricted model



Refining the model

- re-estimated SIMMAP counts

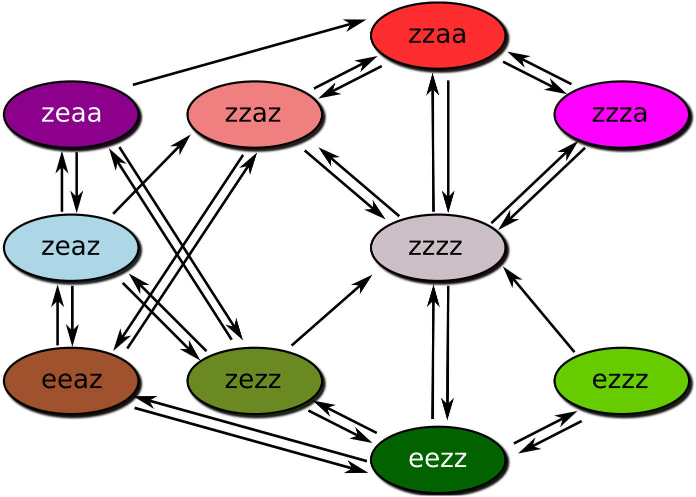
	zzzz	zzaa	zzaz	eezz	zeaz	eeaz	ezzz	zezz	zeaa	zzza
zzzz	-	11.0	31.2	17.0	0.0	0.0	0.8	1.2	0.0	0.8
zzaa	11.5	-	7.2	0.0	0.0	0.0	0.0	0.0	0.7	0.6
zzaz	18.1	11.3	-	0.0	1.0	3.8	0.0	0.0	0.0	0.0
eezz	9.9	0.0	0.0	-	0.0	11.3	1.0	8.2	0.0	0.0
zeaz	0.0	0.0	1.1	0.0	-	2.3	0.0	6.1	4.2	0.0
eeaz	0.0	0.0	5.3	10.5	4.6	-	0.0	0.0	0.0	0.0
ezzz	1.2	0.0	0.0	1.2	0.0	0.0	-	0.0	0.0	0.0
zezz	1.3	0.0	0.0	9.0	4.3	0.0	0.0	-	1.1	0.0
zeaa	0.0	2.5	0.0	0.0	1.6	0.0	0.0	1.0	-	0.0
zzza	0.9	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-

Refining the model

- numbers suggested that transition *no subect marking* → *differential subject marking* is highly unlikeliy

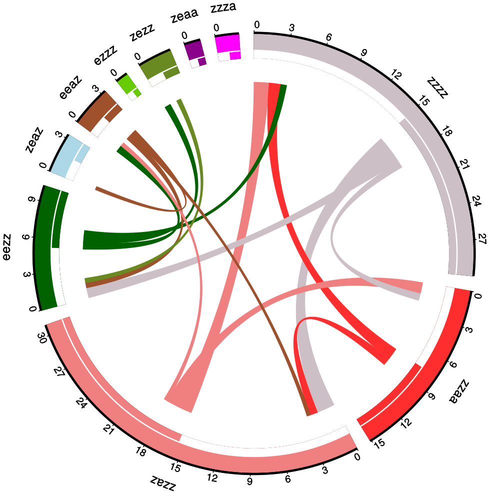
	zzzz	zzaa	zzaz	eezz	zeaz	eeaz	ezzz	zezz	zeaa	zzza
zzzz	-	11.0	31.2	17.0	0.0	0.0	0.8	1.2	0.0	0.8
zzaa	11.5	-	7.2	0.0	0.0	0.0	0.0	0.0	0.7	0.6
zzaz	18.1	11.3	-	0.0	1.0	3.8	0.0	0.0	0.0	0.0
eezz	9.9	0.0	0.0	-	0.0	11.3	1.0	8.2	0.0	0.0
zeaz	0.0	0.0	1.1	0.0	-	2.3	0.0	6.1	4.2	0.0
eeaz	0.0	0.0	5.3	10.5	4.6	-	0.0	0.0	0.0	0.0
ezzz	1.2	0.0	0.0	1.2	0.0	0.0	-	0.0	0.0	0.0
zezz	1.3	0.0	0.0	9.0	4.3	0.0	0.0	-	1.1	0.0
zeaa	0.0	2.5	0.0	0.0	1.6	0.0	0.0	1.0	-	0.0
zzza	0.9	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-

Refining the model

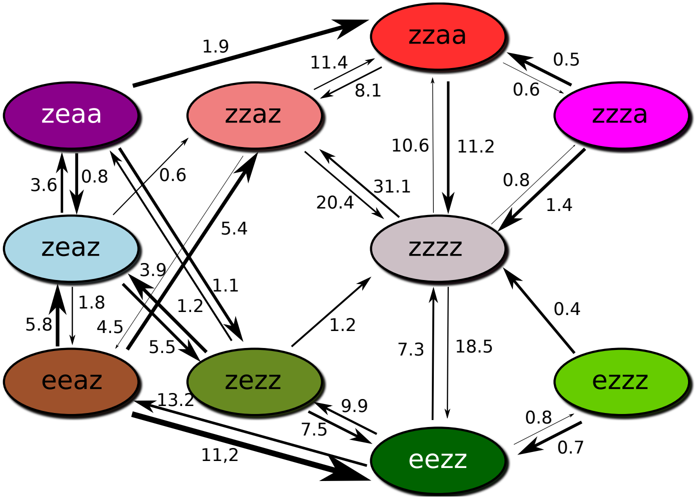


Refining the model

- further re-estimation of parameters with further restricted model
- $\Delta \text{AICM} = 8.7$ in favor of smaller model

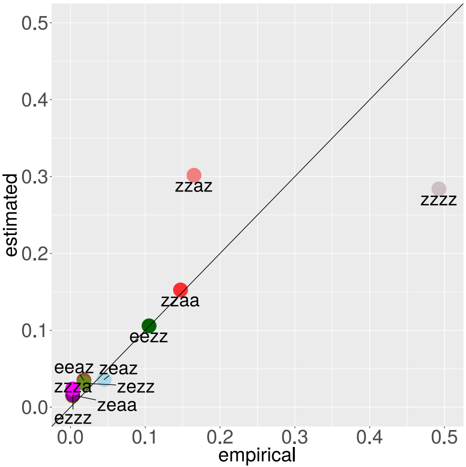


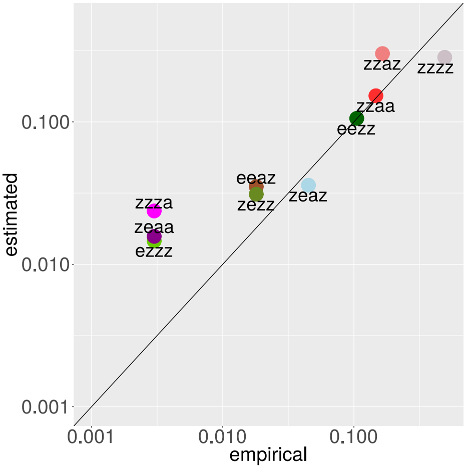
The fitted model



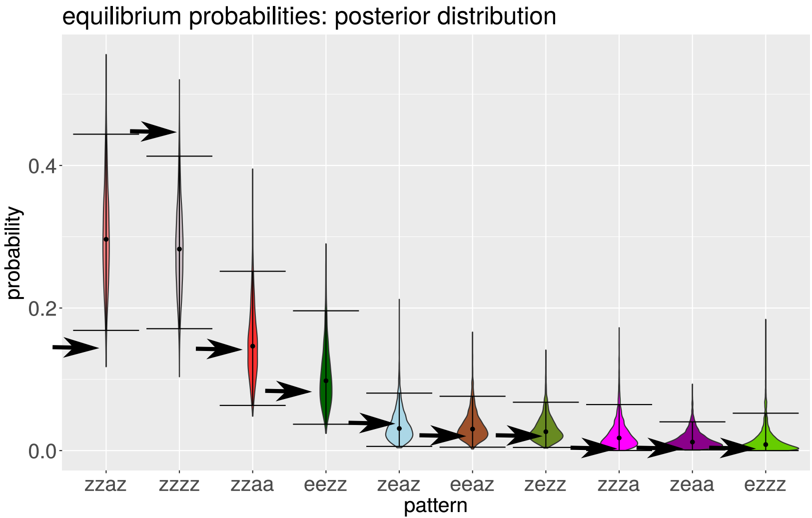
Equilibrium probabilities

Empirical vs. estimated percentages





Equilibrium probabilities



Intermediate summary

- only four patterns occur with more than 10% equilibrium probability:
  - differential object marking
  - no case marking
  - non-differential object marking
  - non-differential subject marking
- patterns inconsistent with Silverstein hierarchy only have joint equilibrium probability of 4% (HDI: [0.3%, 9.4%])
- ergative systems (including mixed ergative/accusative systems) have equilibrium probability of 23% (HDI: [10%, 39%])

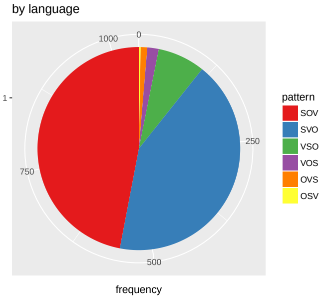
Major word orders

Statistics of major word order distribution

- data: WALS intersected with ASJP
- 1,045 languages, 211 lineages, 32 families with at least 5 languages

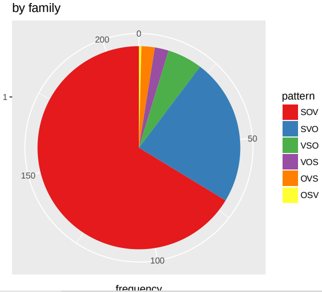
Raw numbers

SOV	SVO	VSO	VOS	OVS	OSV
491	442	79	19	11	3
47.0%	42.3%	7.6%	1.8%	1.1%	0.3%



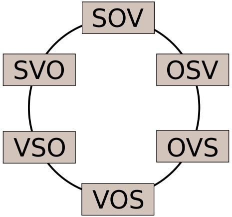
Weighted by lineages

SOV	SVO	VSO	VOS	OVS	OSV
139.1	49.3	11.8	4.7	4.5	0.8
66.3%	23.4%	5.6%	2.2%	2.1%	0.4%



Previous approaches

- Gell-Mann and Ruhlen (2011):
  - Proto-world was SOV
  - general pathway: SOV $\rightarrow$ SVO $\leftrightarrow$ VSO/VOS
  - minor pathway: SOV $\rightarrow$ OVS/OSV
  - exceptions due to diffusion
- Ferrer-i-Cancho (2015):



  - permutation circle
  - transition probability inversely related to path length

Previous approaches

- Maurits and Griffiths (2014):
  - Bayesian rate estimation, based on five families and NJ-trees

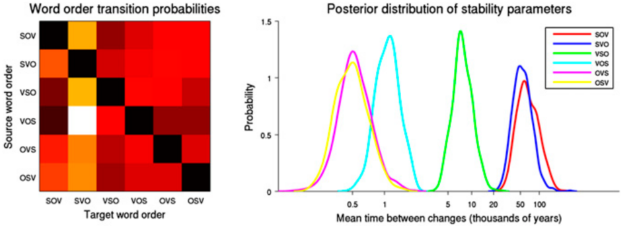


Fig. 1. Results of inferring a single mutation matrix Q for all six language families. (*Left*) Heat map showing the transition probabilities between word orders. Higher intensity (white, yellow) indicates more-probable transitions compared with lower intensity (red, brown), so SOV is most likely to transition to SVO and SVO to SOV. VSO is much more likely to transition to SVO than to SOV. (*Right*) Inferred posterior distributions of stability parameters for each word order. The horizontal axis shows the stability parameter, expressed as the mean time between transitions; i.e., higher values indicate a more stable word order.

Workflow

(data from all 32 families with ≥ 5 languages in data base)

- estimate posterior tree distributions with MrBayes for each family, using Glottolog as constraint tree
- test whether universal or lineage-specific model gives a better fit
- estimate transition rates with best model
- estimate stationary distribution of major word order categories
- apply *stochastic character mapping* (SIMMAP; Bollback 2006)
- estimate expected number of mutations for each transition type

Estimating posterior tree distributions

- using characters extracted from ASJP data (Jäger 2018)
- Glottolog as constraint tree
- Γ-distributed rates
- ascertainment bias correction
- relaxed molecular clock (IGR)
- uniform tree prior
- stop rule: 0.01, samplefreq=1000
- if convergence later than after 1,000,000 steps, sample 1,000 trees from posterior

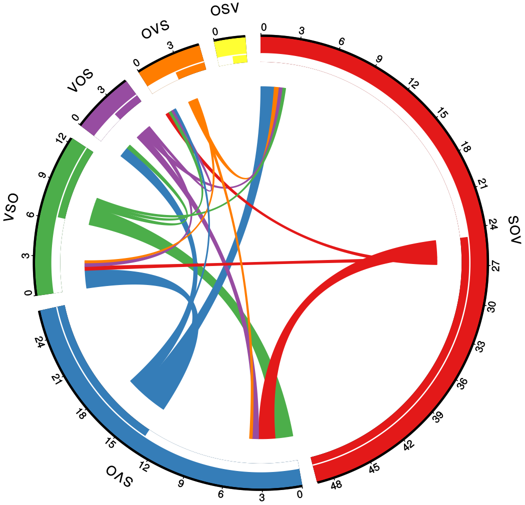
Rate estimation

Unrestricted model

	SOV	SVO	VSO	VOS	OVS	OSV
SOV	—	r_1	r_2	r_3	r_4	r_5
SVO	r_6	—	r_7	r_8	r_9	r_{10}
VSO	r_{11}	r_{12}	—	r_{13}	r_{14}	r_{15}
VOS	r_{16}	r_{17}	r_{18}	—	r_{19}	r_{20}
OVS	r_{21}	r_{22}	r_{23}	r_{24}	—	r_{25}
OSV	r_{26}	r_{27}	r_{28}	r_{29}	r_{30}	—

- Fitting a separate unrestricted model for each family is preferred with a Bayes Factor 120, so there is actually little evidence for a universal dynamics
- in line with the literature, I will nevertheless assume that all lineages have the same underlying dynamics

Fitted rates



Rate estimation

GTR model

	SOV	SVO	VSO	VOS	OVS	OSV
SOV	—	$r_1\pi_2$	$r_2\pi_3$	$r_3\pi_4$	$r_4\pi_5$	$r_5\pi_6$
SVO	$r_1\pi_1$	—	$r_6\pi_3$	$r_7\pi_4$	$r_8\pi_5$	$r_9\pi_6$
VSO	$r_2\pi_1$	$r_6\pi_2$	—	$r_{10}\pi_4$	$r_{11}\pi_5$	$r_{12}\pi_6$
VOS	$r_3\pi_1$	$r_7\pi_2$	$r_{10}\pi_3$	—	$r_{13}\pi_5$	$r_{14}\pi_6$
OVS	$r_4\pi_1$	$r_8\pi_2$	$r_{11}\pi_3$	$r_{13}\pi_4$	—	$r_{15}\pi_6$
OSV	$r_5\pi_1$	$r_9\pi_2$	$r_{12}\pi_3$	$r_{14}\pi_4$	$r_{15}\pi_5$	—

- Bayes factor in favor of GTR model: 13

Rate estimation

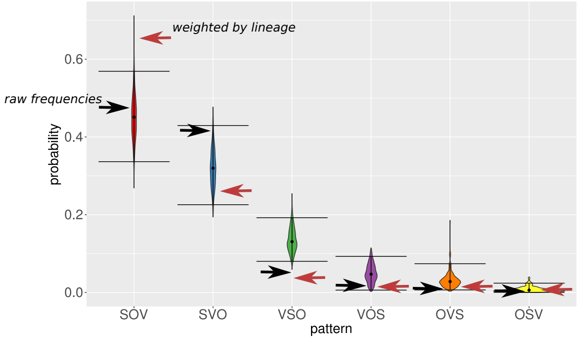
GTR model: fitted values (mean posterior) per 1,000 units of time

	SOV	SVO	VSO	VOS	OVS	OSV
SOV	−6.67	4.14	0.63	0.29	1.32	0.29
SVO	5.83	−19.18	8.03	4.02	0.88	0.43
VSO	2.09	19.05	−25.75	2.91	1.33	0.37
VOS	2.64	26.10	7.95	−38.48	1.40	0.39
OVS	18.15	8.58	5.49	2.11	−34.71	0.39
OSV	15.71	16.68	5.95	2.11	1.53	−41.99

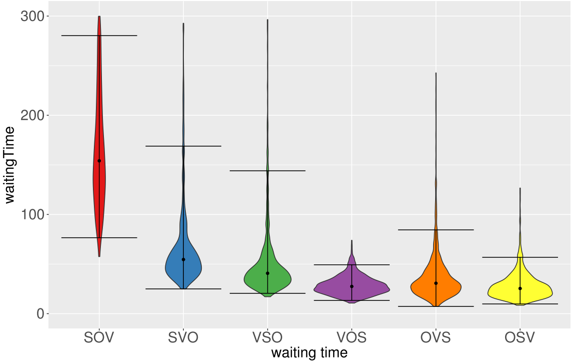
SOV	SVO	VSO	VOS	OVS	OSV
45.2%	32.2%	13.6%	5.0%	3.3%	0.8%

Rate estimation

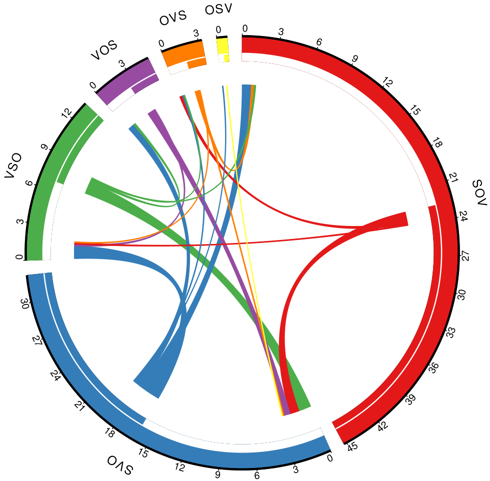
Equilibrium probabilities (posterior distribution)



Stability (posterior distribution)



Rate estimation



Reconstruction history with SIMMAP

- estimated frequency of mutations within the 32 families under consideration (posterior mean, 100 iterations)

	SOV	SVO	VSO	VOS	OVS	OSV
SOV	–	16.55	2.51	0.15	3.26	0.04
SVO	15.89	–	17.28	11.34	1.31	1.02
VSO	1.29	14.11	–	2.02	1.03	0.03
VOS	0.11	5.22	1.54	–	0.06	0.00
OVS	2.33	0.63	0.28	0.01	–	0.03
OSV	0.07	0.27	0.04	0.00	0.00	–

Inconsistent with Ramon-i-Ferrer's model

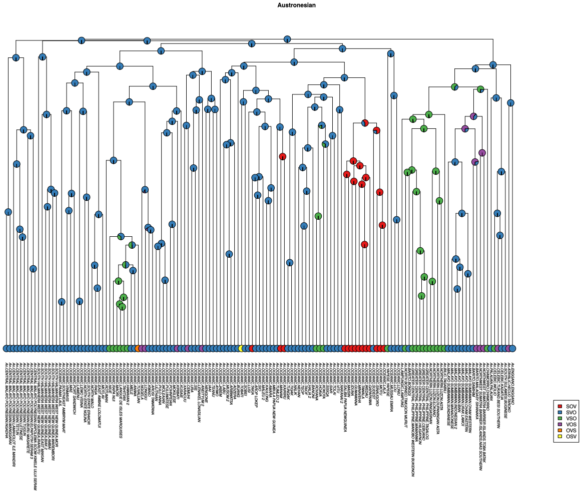
	SOV	SVO	VSO	VOS	OVS	OSV
SOV	–	16.55	2.51	0.15	3.26	0.04
SVO	15.89	–	17.28	11.34	1.31	1.02
VSO	1.29	14.11	–	2.02	1.03	0.03
VOS	0.11	5.22	1.54	–	0.06	0.00
OVS	2.33	0.63	0.28	0.01	–	0.03
OSV	0.07	0.27	0.04	0.00	0.00	–

Inconsistent with Gell-Mann and Ruhlen's model

	SOV	SVO	VSO	VOS	OVS	OSV
SOV	–	16.55	2.51	0.15	3.26	0.04
SVO	15.89	–	17.28	11.34	1.31	1.02
VSO	1.29	14.11	–	2.02	1.03	0.03
VOS	0.11	5.22	1.54	–	0.06	0.00
OVS	2.33	0.63	0.28	0.01	–	0.03
OSV	0.07	0.27	0.04	0.00	0.00	–

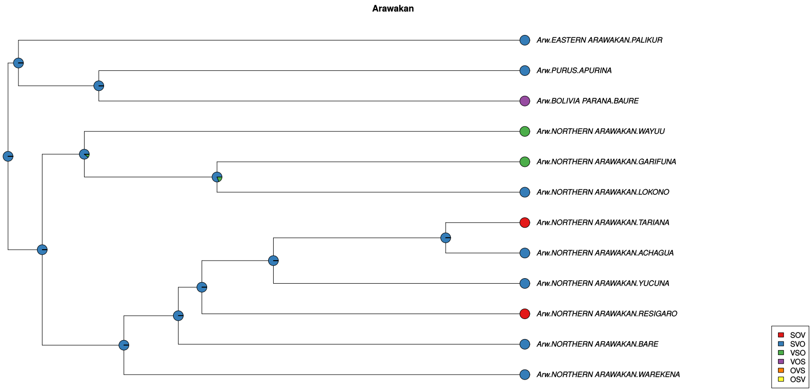
Counterexamples to Ramon-i-Ferrer's model: SVO $\rightarrow$ VOS

Austronesian: 8 such mutations expected



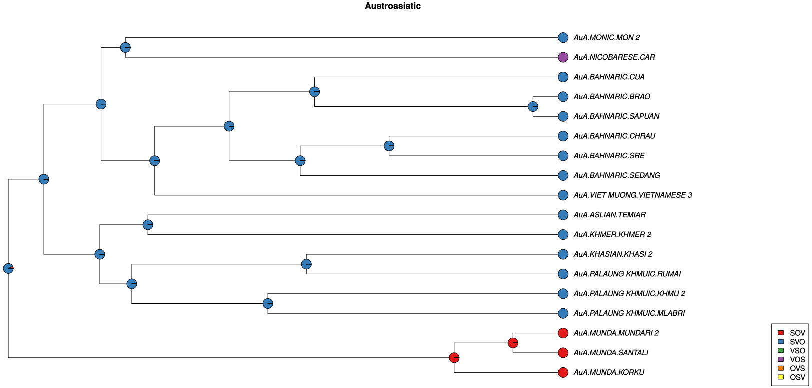
Counterexamples to Ramon-i-Ferrer's model: SVO → VOS

Arawakan: 1 such mutation expected



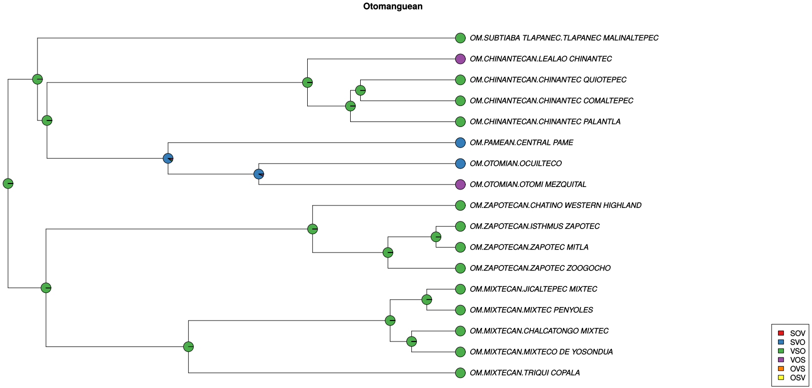
Counterexamples to Ramon-i-Ferrer's model: SVO → VOS

Austroasiatic: 1 such mutation expected



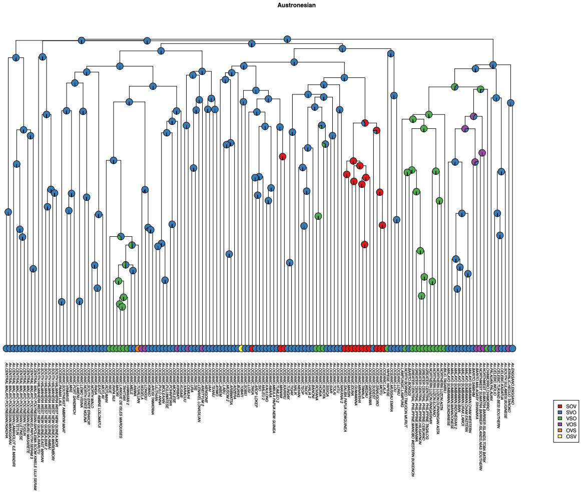
Counterexamples to Ramon-i-Ferrer's model: SVO → VOS

Otomanguean: 1 such mutation expected



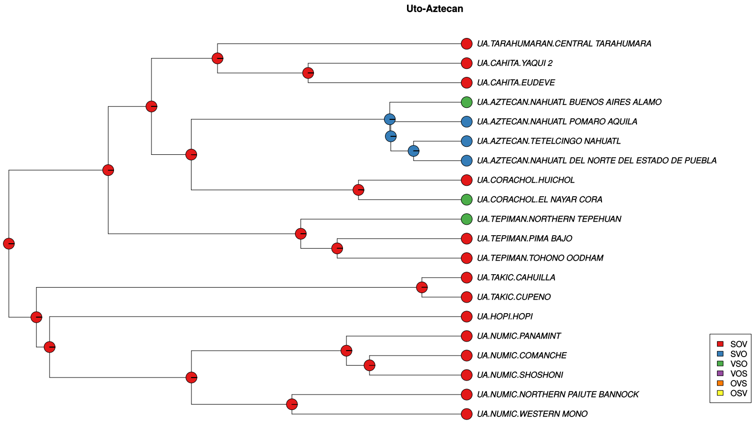
Counterexamples to Ramon-i-Ferrer's model: VOS $\rightarrow$ SVO

3.9 mutations such expected in Austronesian



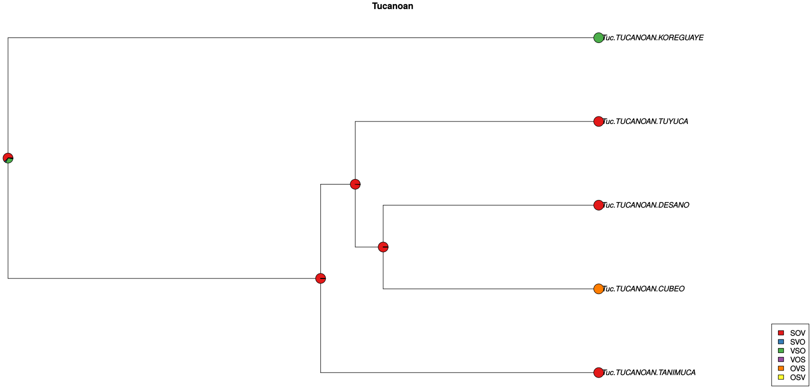
Counterexamples to Ramon-i-Ferrer's model: SOV $\rightarrow$ VSO

Uto-Aztecan: 1.8 such mutatations expected



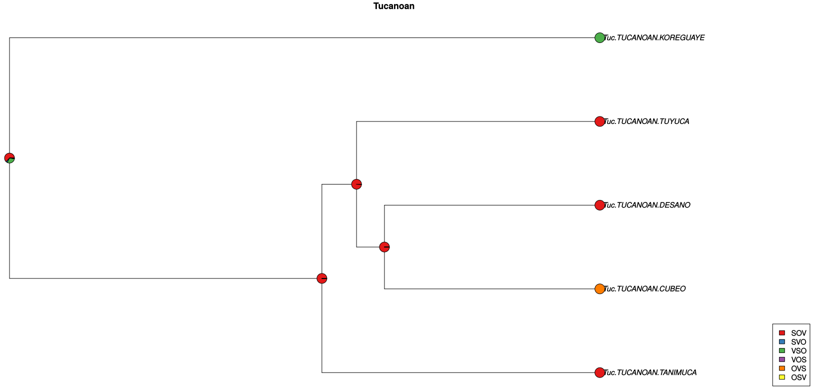
Counterexamples to Ramon-i-Ferrer's model: SOV → VSO

Tucanoan: 0.5 such mutatations expected



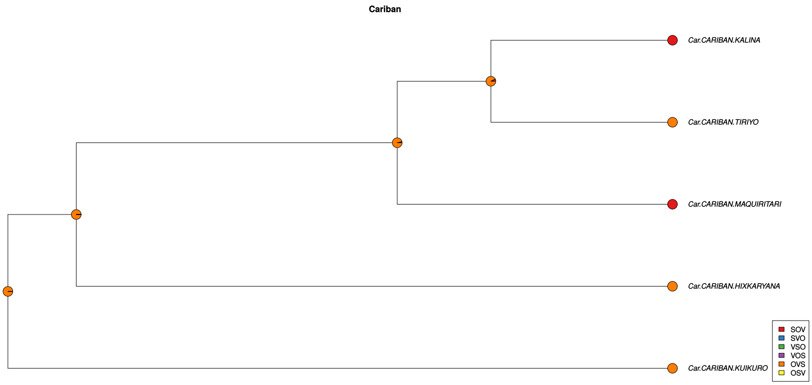
Counterexamples to Ramon-i-Ferrer's model: SOV $\rightarrow$ OVS

Tucanoan: 1 such mutatations expected



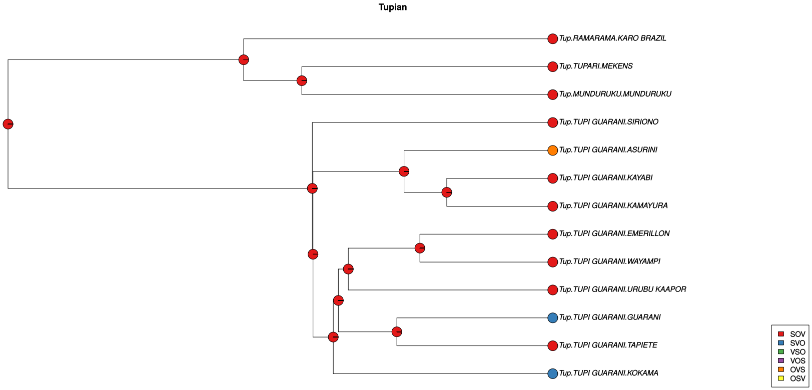
Counterexamples to Ramon-i-Ferrer's model: SOV $\rightarrow$ OVS

Cariban: 1 such mutatations expected



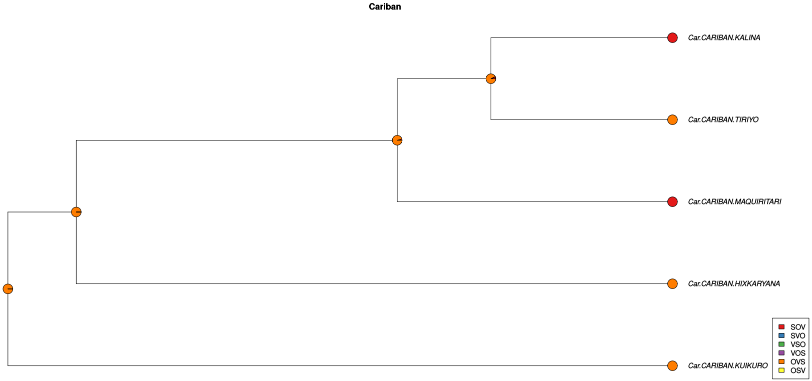
Counterexamples to Ramon-i-Ferrer's model: SOV → OVS

Tupian: 1 such mutatations expected



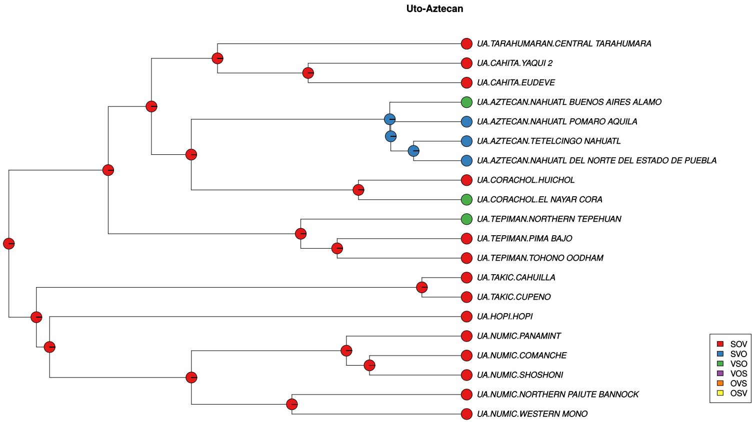
Counterexamples to Ramon-i-Ferrer's model: OVS $\rightarrow$ SOV

Cariban: 1.3 such mutatations expected



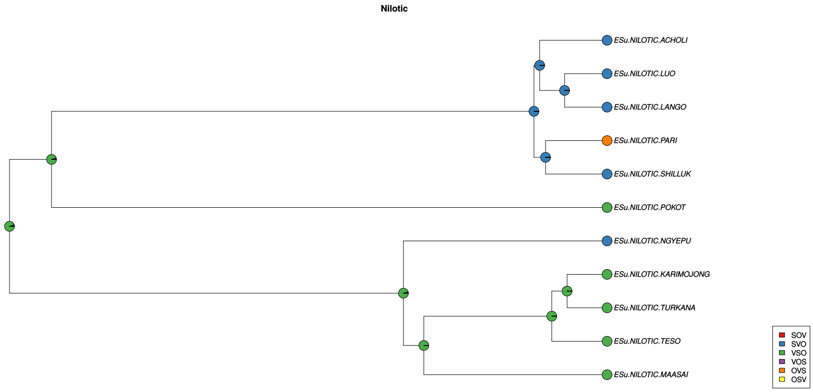
Counterexamples to Ramon-i-Ferrer's model: VSO → SOV

Uto-Aztecan: 0.8 such mutatations expected



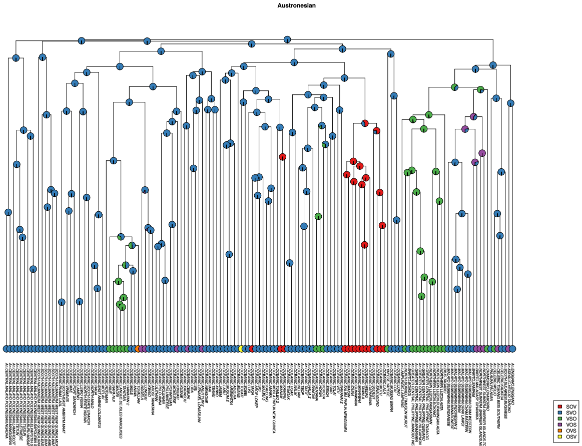
Counterexamples to Ramon-i-Ferrer's model: SVO → OVS

Nilotic: 0.9 such mutatation expected



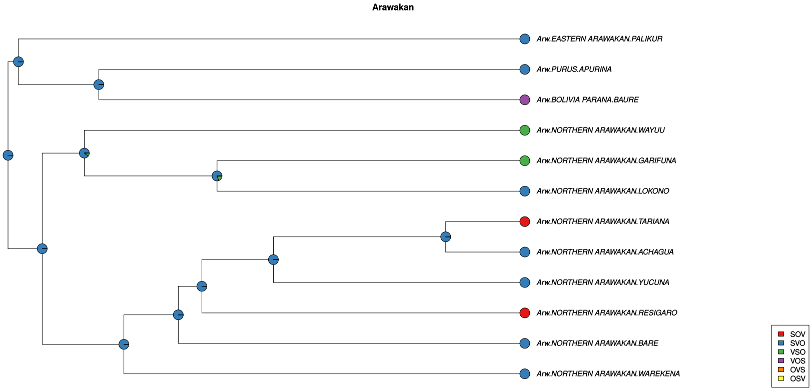
Additional counterexamples to Gell-Mann & Ruhlen's model: SVO $\rightarrow$ SOV

Austronesian: 3 such mutatations expected



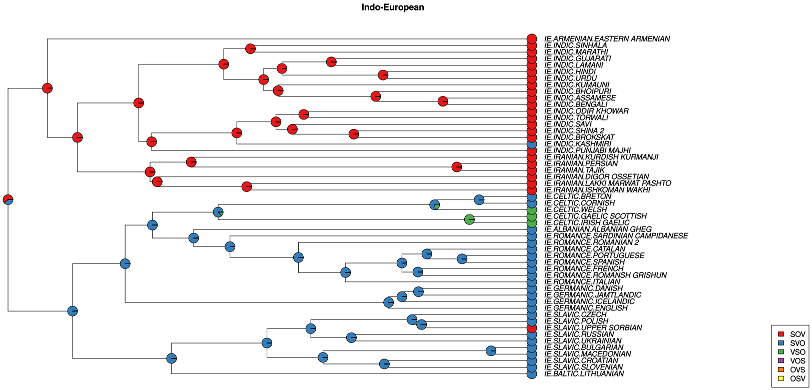
Additional counterexamples to Gell-Mann & Ruhlen's model: SVO $\rightarrow$ SOV

Arawakan: 2 such mutatations expected



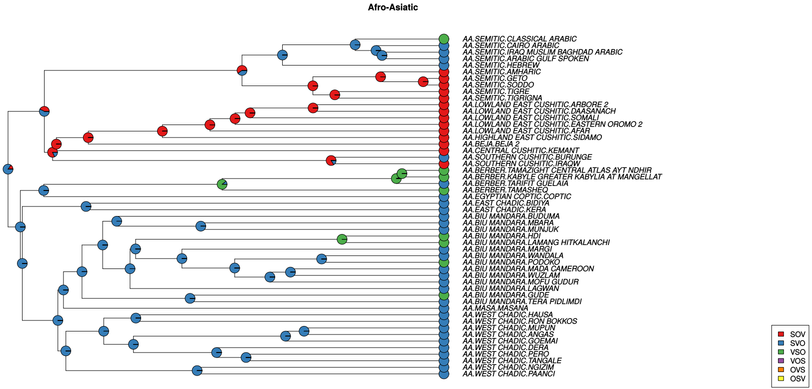
Additional counterexamples to Gell-Mann & Ruhlen's model: SVO $\rightarrow$ SOV

Indo-European: 2 such mutatations expected



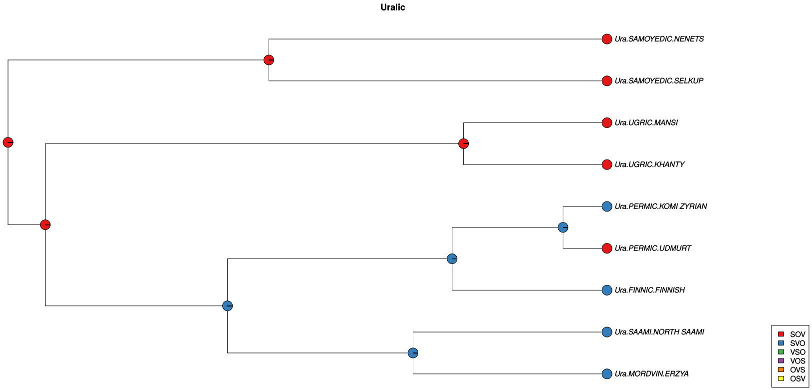
Additional counterexamples to Gell-Mann & Ruhlen's model: SVO → SOV

Afro-Asiatic: 1.6 such mutatations expected



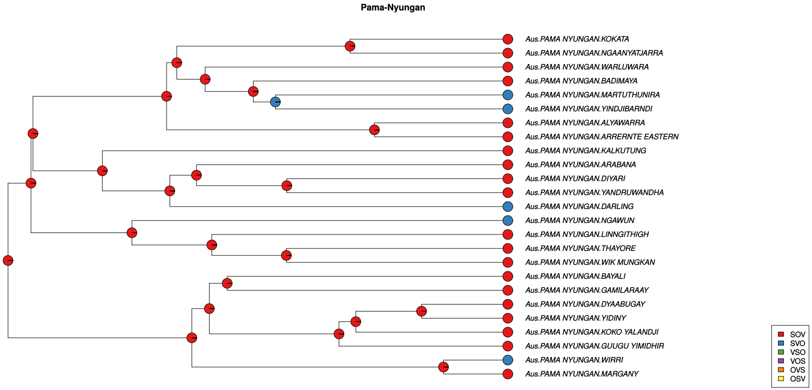
Additional counterexamples to Gell-Mann & Ruhlen's model: SVO $\rightarrow$ SOV

Uralic: 1.4 such mutatations expected



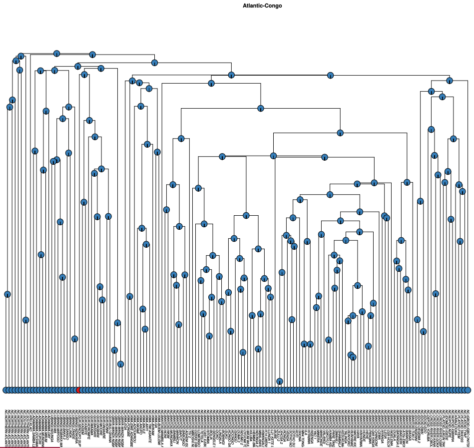
Additional counterexamples to Gell-Mann & Ruhlen's model: SVO $\rightarrow$ SOV

Pama-Nyungan: 1.2 such mutatations expected



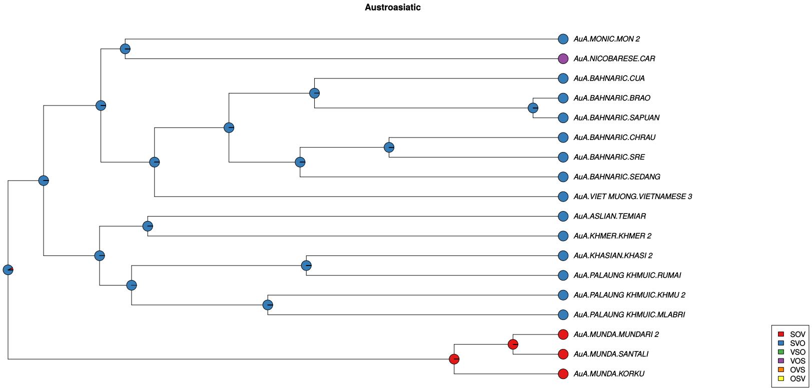
Additional counterexamples to Gell-Mann & Ruhlen's model: SVO $\rightarrow$ SOV

Atlantic-Congo: 1 such mutatations expected



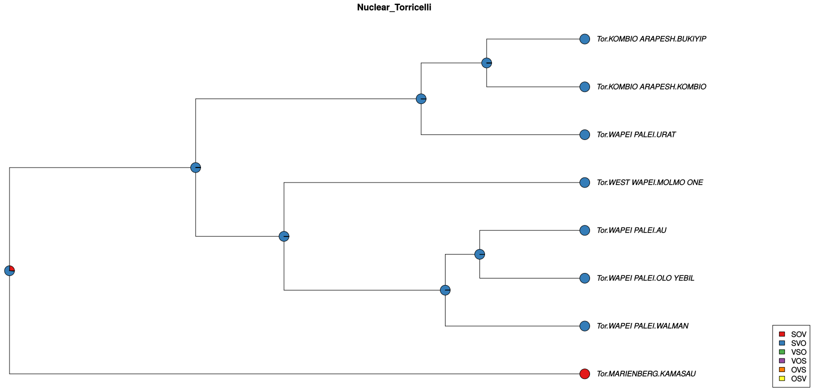
Additional counterexamples to Gell-Mann & Ruhlen's model: SVO $\rightarrow$ SOV

Austroasiatic: 0.9 such mutatations expected



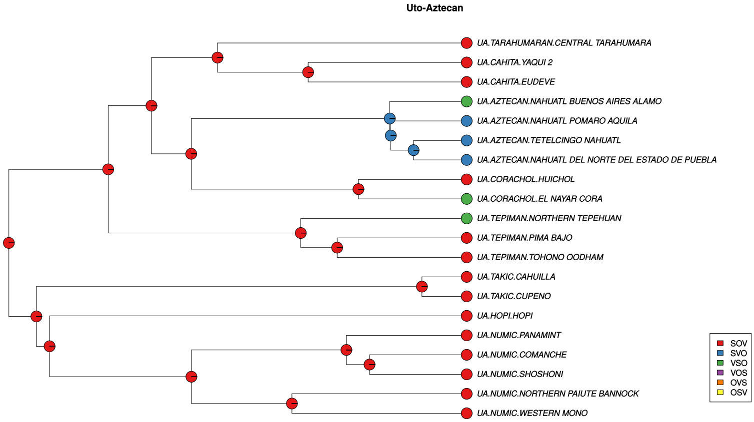
Additional counterexamples to Gell-Mann & Ruhlen's model: SVO $\rightarrow$ SOV

Nuclear Toricelli: 0.9 such mutatations expected



Additional counterexamples to Gell-Mann & Ruhlen's model: SOV $\rightarrow$ VSO

Uto-Aztecan: 1.8 such mutatations expected



Summary

- no evidence for general preference of SOV $\rightarrow$ SVO over the reverse
- SVO is currently over-represented due to recent spread of Austronesian and Atlantic-Congo, but not excessively so
- multiple counter-evidence to Ramon-i-Ferrer's and Gell-Mann & Ruhlen's models

Judith Aissen. Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory*, 21(3):435–483, 2003.

Balthasar Bickel, Alena Witzlack-Makarevich, and Taras Zakharko. Typological evidence against universal effects of referential scales on case alignment. In Ina Bornkessel-Schlesewsky, Andrej L. Malchukov, and Marc D. Richards, editors, *Scales and hierarchies: A cross-disciplinary perspective*, pages 7–43. de Gruyter, Berlin/Munich/Boston, 2015.

Jonathan P. Bollback. SIMMAP: stochastic character mapping of discrete traits on phylogenies. *BMC Bioinformatics*, 7(1):88, 2006.

Georg Bossong. *Differentielle Objektmarkierung in den neuiranischen Sprachen*. Günther Narr Verlag, Tübingen, 1985.

Bernard Comrie. *Language Universals and Linguistic Typology*. Basil Blackwell, Oxford, 1981.

Ramon Ferrer-i-Cancho. Kauffman's adjacent possible in word order evolution. arXiv preprint arXiv:1512.05582, 2015.

Murray Gell-Mann and Merritt Ruhlen. The origin and evolution of word order. *Proceedings of the National Academy of Sciences*, 108(42):17290–17295, 2011.

Sebastian Höhna, Michael J. Landis, Tracy A. Heath, Bastien Boussau, Nicolas Lartillot, Brian R. Moore, John P. Huelsenbeck, and Frederik Ronquist. Revbayes: Bayesian phylogenetic inference using graphical models and an interactive model-specification language. *Systematic biology*, 65(4):726–736, 2016.

Gerhard Jäger. Phylogenetic inference from word lists using weighted alignment with empirically determined weights. *Language Dynamics and Change*, 3(2): 245–291, 2013.

Gerhard Jäger. Support for linguistic macrofamilies from weighted sequence alignment. *Proceedings of the National Academy of Sciences*, 112(41):12752–12757, 2015. doi: 10.1073/pnas.1500331112.

Gerhard Jäger. Global-scale phylogenetic linguistic inference from lexical resources. arXiv:1802.06079, 2018.

Gerhard Jäger and Søren Wichmann. Inferring the world tree of languages from word lists. In S. G. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Feher, and T. Verhoef, editors, *The Evolution of Language: Proceedings of the 11th International Conference (EVOLANG11)*, 2016. Available online: http://evolang.org/neworleans/papers/147.html.

Elena Maslova. A dynamic approach to the verification of distributional universals. *Linguistic Typology*, 4(3):307–333, 2000.

Luke Maurits and Thomas L. Griffiths. Tracing the roots of syntax with Bayesian phylogenetics. *Proceedings of the National Academy of Sciences*, 111(37): 13576–13581, 2014.

Mark Pagel and Andrew Meade. Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. *The American Naturalist*, 167(6):808–825, 2006.

Mark Pagel and Andrew Meade. BayesTraits 2.0. software distributed by the authors, November 2014.

Hugo Reyes-Centeno, Katerina Harvati, and Gerhard Jäger. Tracking modern human population history from linguistic and cranial phenotype. *Scientific Reports*, 6, 2016.

Frederik Ronquist and John P. Huelsenbeck. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19(12):1572–1574, 2003.

Michael Silverstein. Hierarchy of features and ergativity. In R. M. W. Dixon, editor, *Grammatical Categories in Australian Languages*, pages 112–171. Australian Institute of Aboriginal Studies, Canberra, 1976.

Søren Wichmann, Eric W. Holman, and Cecil H. Brown. The ASJP database (version 17). http://asjp.clld.org/, 2016.