Understanding Language Evolution Using an Event-Based Model

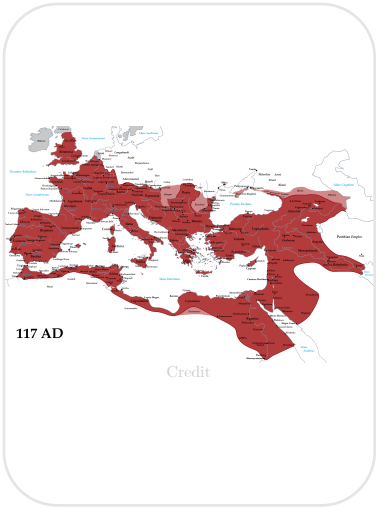
John Huelsenbeck
Department of Integrative Biology
johnh@berkeley.edu

Introduction

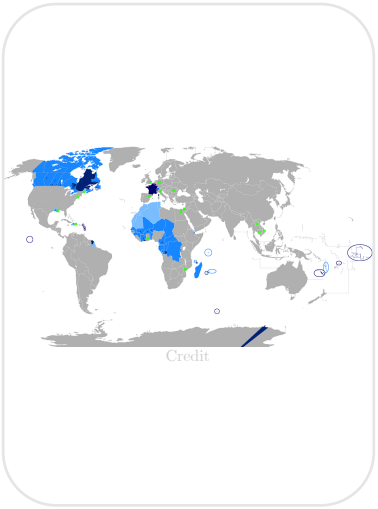
In this paper, we attempt to do the impossible!

Languages

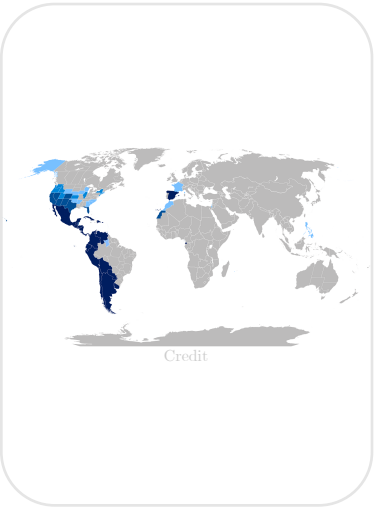
Latin



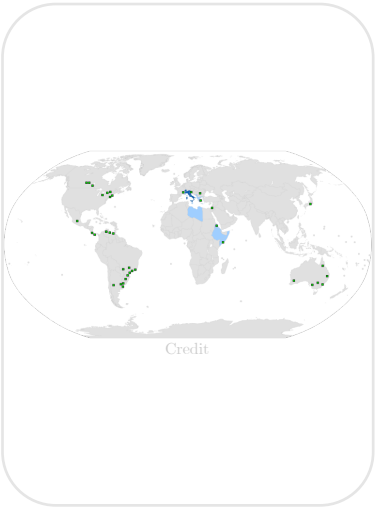
French



Spanish



Italian

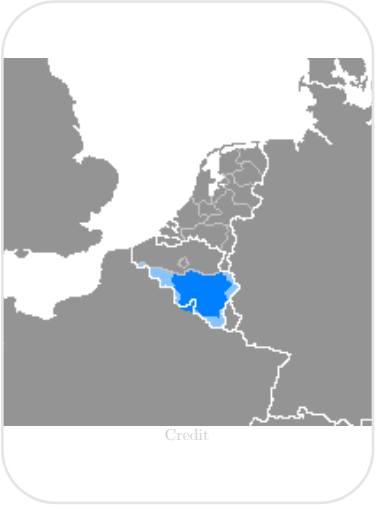


Brazilian Portuguese

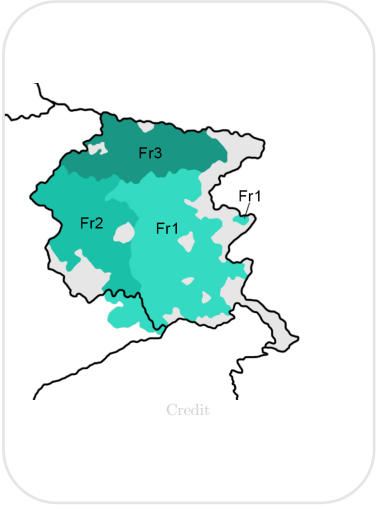
Portuguese

Catalan

Walloon



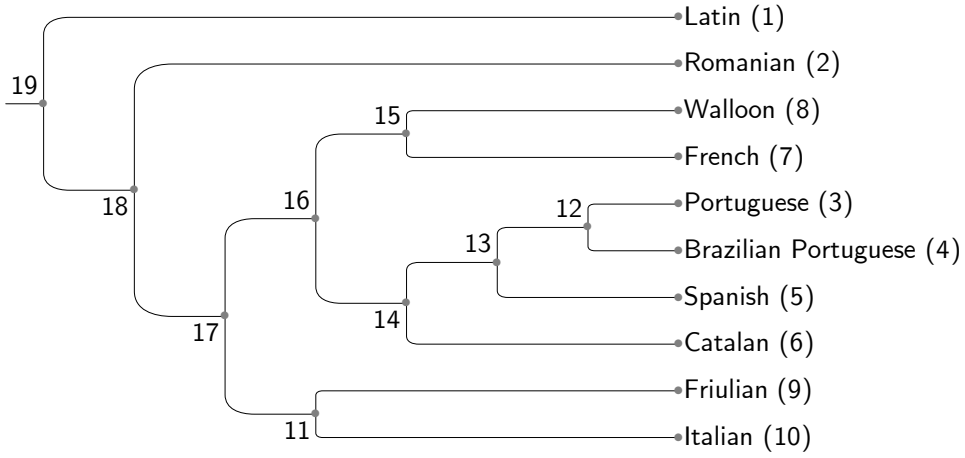
Friulian



Romanian

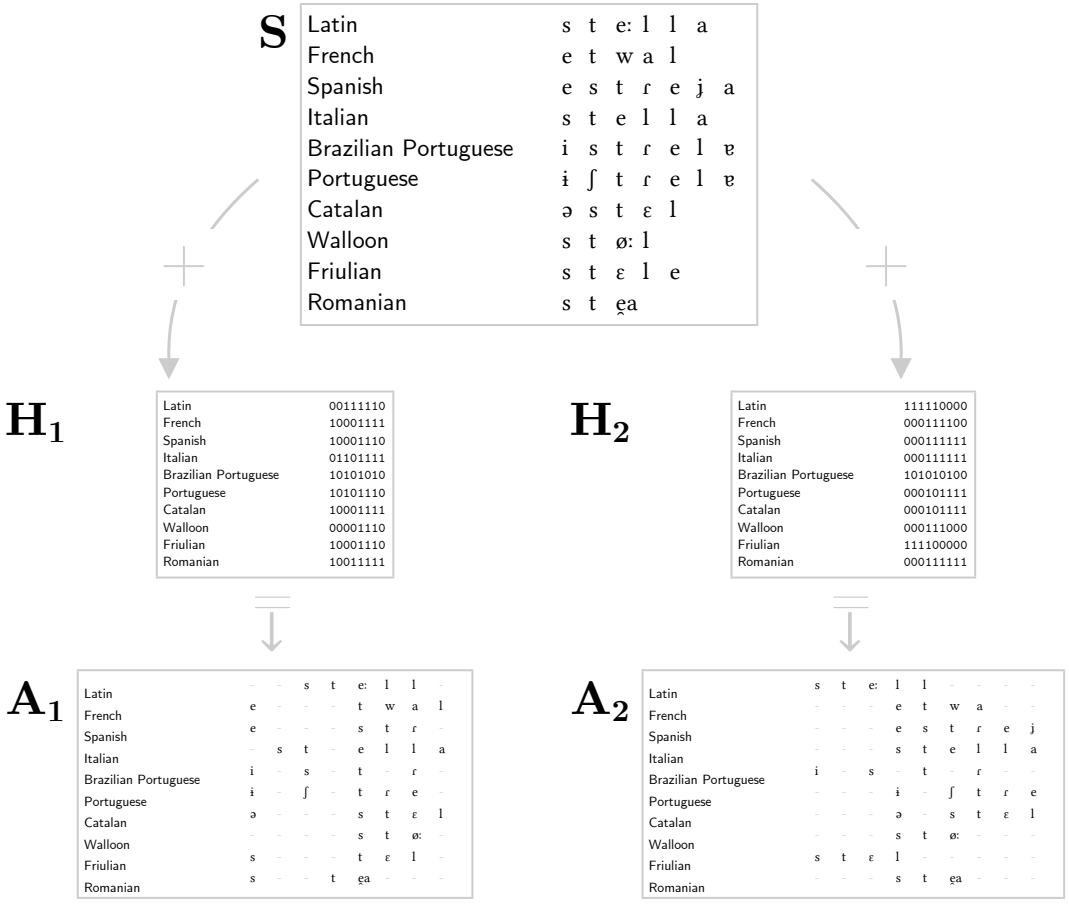
Example Tree

An example tree showing the relationships of $N = 10$ languages.



Alignment

Alignments (**A**) are formed from the observed segments (**S**) and a homology map (**H**).



Character Assignments

Each segment gets a different number

0	e	1	g	2	oː	3	ʒ	4	ə	5	j̊	6	o	7	i	8	w	9	ɔ	10	ʤ	11	ɪ	12	j	13	n	14	s
15	u	16	ʃ	17	oj	18	v	19	b	20	t	21	uː	22	y	23	d	24	f	25	eː	26	m	27	a				
28	ɾ	29	ɐ	30	œ	31	ɛ	32	l	33	r	34	x	35	ʎ	36	ʁ	37	iː	38	ɲ	39	o̥a	40	h				
41	ẽj̃	42	ɔ̃	43	ẽ	44	ɑ̃	45	ĩ	46	ɐ̃	47	ʧ	48	ɨ	49	ã	50	p	51	z	52	ð	53	k	54	ʦ		
55	aː	56	ɐ̃j̃	57	ʁ	58	ɒ	59	θ	60	ej	61	ɐj	62	õ	63	ɛː	64	ɥ	65	øː	66	ɬ	67	c				
68	ɛj	69	ø	70	e̥a	71	kʷ	72	ɣ	73	aj	74	g	75	ŋ	76	gʷ	77	ẽu̥	78	ɑ	79	β	80	ɟ				
81	ɛ̃	82	ɔː	83	ũ	84	j̃	85	œ̃	86	au̥	87	ʧ	88	w̃														

Partition Assignments

Linguistic Model 1

1 Nasal Vowel

ẽj̃ ɔ̃ ẽ ɑ̃ ĩ ɐ̃ ã ɐ̃j̃ õ ẽu̯ ɛ̃ ũ œ̃

2 Vowel

e oː ə o i ɔ ɪ u oj uː y eː a ɐ œ ɛ iː o̯a ɨ aː ɒ ej ɐj ɛː øː ɛj ø e̯a aj ɑ ɔː au̯

3 Nasal Consonant

n m ɲ ŋ j̃ w̃

4 Non Sylabic Sonorants

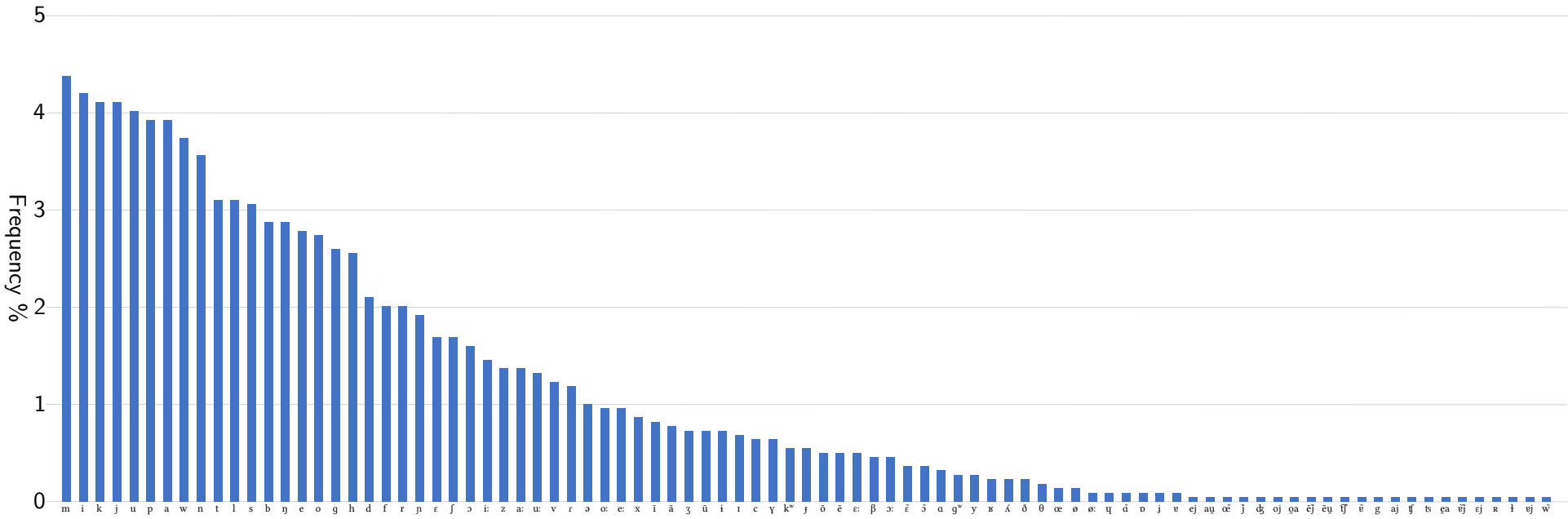
w j l r

5 Consonants

ɡ ʒ ʝ ʤ s ʃ v b t d f ɾ x ʎ ʁ h ʧ p z ð k ʦ ʀ θ ɥ ɬ c kʷ ɣ g gʷ β ɟ ʧ

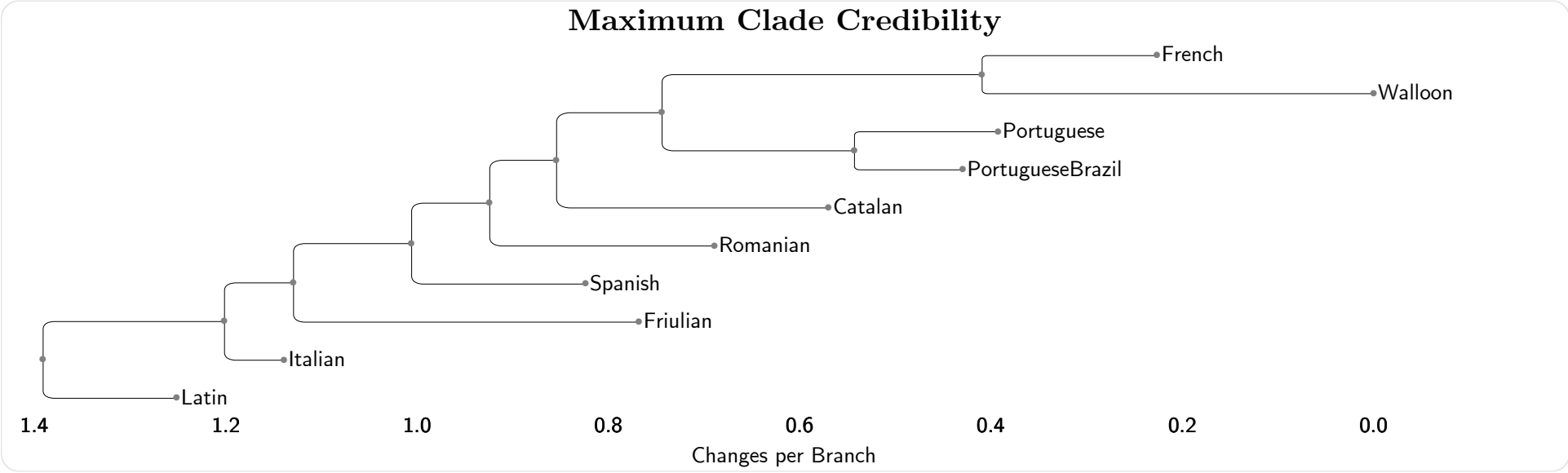
Prior Segment Frequencies

Frequency of occurance of segments in the lexicon [1]



[1] Max Planck Institute for the Science of Human History *PHOIBLE 2.0* ed. Steven Moran, and Daniel McCloy (: Max Planck Institute for the Science of Human History 2019)www.phoible.org,

Results


Maximum Clade Credibility
French
Walloon
Portuguese
PortugueseBrazil
Catalan
Romanian
Spanish
Friulian
Italian
Latin
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Changes per Branch

Questions