Understanding Language Evolution Using an Event-Based Model

David Goldstein
Department of Linguistics
dgoldstein@humnet.ucla.edu

Introduction

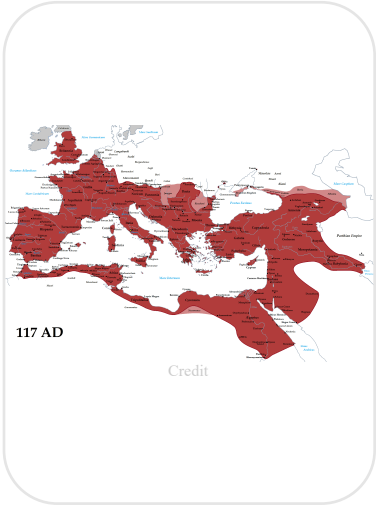
Modern languages are related to one another through a complicated history of divergence and word borrowing. The divergence of languages is caused by the slow change in spoken language as it is passed from parents to offspring. Over time, divergence causes languages to become increasingly different from one another, ultimately to the point where they are mutually unintelligible. Languages that were spoken by the same human group more recently in time are considered to be more closely related to each other than they are to groups that spoke the language more distantly in time; this relatedness information can be depicted by a tree-like diagram called a 'phylogeny.' Linguistic borrowing, by contrast, causes languages to become more similar to one another.

Language	IPA	Coding
English	/hænd/	0
German	/hant/	0
French	/mɛ̃/	1
Spanish	/mano/	1
Italian	/maːno/	1
Russian	/rʊ̟ka/	2
Polish	/rɛŋka/	2

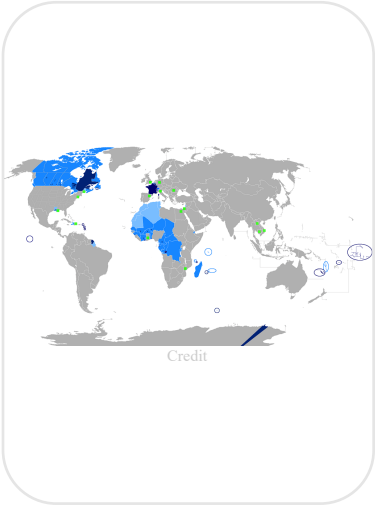
Table 1. Coding of lexical cognates for the word *hand*.

Languages

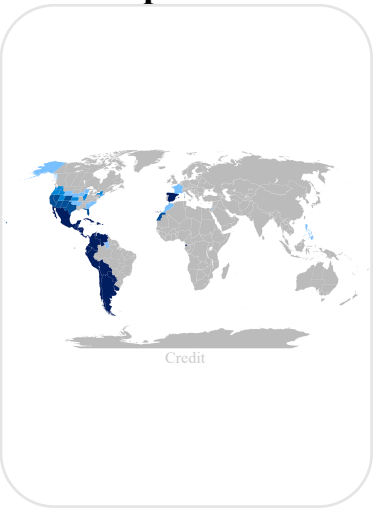
Latin



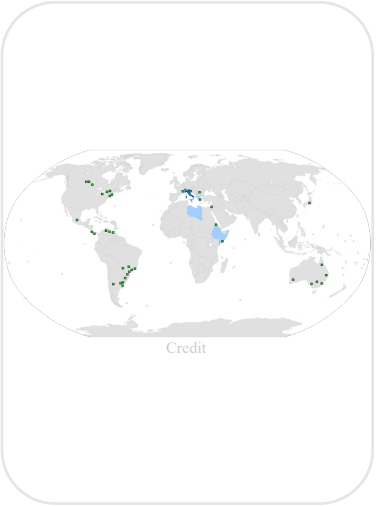
French



Spanish



Italian

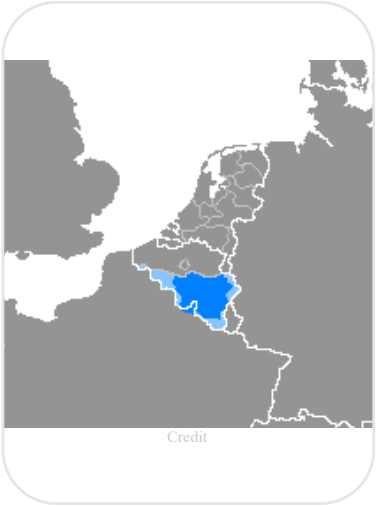


Brazilian Portuguese

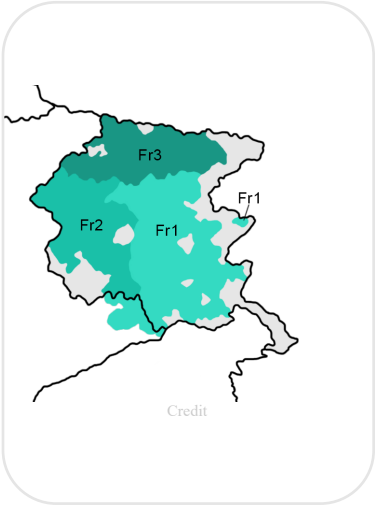
Portuguese

Catalan

Walloon



Friulian



Romanian

Concept Examples

	Child-Primary	Child-Ninnus	Fish-Piskem	Fish-Pescion	Bird-Avem	Bird-Aucellus
Latin	i n f a n t e m		p i s k e m		a w e m	- a k - e l l u s
French	ɑ̃ - f ɑ̃ - - - -			p w a s ɔ̃		w a z - o - - - -
Spanish	i n f a n t e -	n i ɲ o	p e θ - - - -		a b e -	
Italian	- - f a n t e -	n i n o	p e ʃ ʃ e -			- u t ʧ ɛ l l o -
PortugueseBrazil	ĩ - f ɐ̃ - ʧ i -	n e n ẽj̃	p ej ʃ - i -		a v i -	
Portuguese	ĩ - f ɐ̃ - t ɨ -	n e n ẽj̃	p ɐj ʃ - ɨ -		a v ɨ -	
Catalan		n e n -	p e ʃ - - -		a w - -	- u s - e - ʎ - -
Walloon	ɛ - f ɑ̃ - - - -			p - ɛ h õ		- u h - ɛ - - - -
Friulian		n i n i	p ɛ s - - -			- u ʧ j ɛ - l - -
Romanian			p e ʃ t e -			

	Bird-Passerem	Dog-Primary	Louse-Primary	Leaf-Primary	Leaf-Frondea	Root-Primary
Latin	p a s s e r e m	k - a n e m	p - eː d u - k u l u s	f - o l i a	f r o n d e a	- r aː d - iː k e m
French		ʃ j ɛ̃ - - -	p - u - - - - - - -	f - œ j - -		
Spanish	p a - x a ɾ o -	k - a n - -	p j o - - - x - - o - - - o x - a			- r a - - i θ - -
Italian	p a s s e r o -	k - a n e -	p - i d ɔ k k - j o -	f - ɔ ʎ ʎ a	f r o n ʤ - a	
PortugueseBrazil	p a - s ɐ ɾ u -	k - ɐ̃ w̃ - - -	p - i - o - - ʎ u - -	f - o ʎ - ɐ	f ɾ õ - s - ɐ	- ʁ a - - i s - -
Portuguese	p a - s ɐ ɾ u -	k - ɐ̃ w̃ - - -	p - i - o - - ʎ u - -	f - o ʎ - ɐ	f ɾ õ - s - ɐ	- ʁ ɐ - - i ʃ - -
Catalan		k - a - - -	p - o - - - - ʎ - - -	f - u ʎ - ə		ə r ɛ l - - - - -
Walloon		ʧ - ɛ̃ - - -	p j u - - - - - - -	f - ɔj - - -		
Friulian	p a - s a r - -	c - a n - -	p - e d oː - - l i - -	f w e - - e		- r a d r iː s - -
Romanian	p a - s ə ɾ e -	k - ɨj n e -	p - ə d u - k - j e	f - o̯a - j e	f ɾ u n z - ə	

	Root-Radicina	Rope-Primary	Rope-Restis	Skin-Primary	Skin-Pellis	Meat-Primary
Latin	r aː d iː k iː n a	f uː n - e m	r e s t - e m		p - e l l e m	k a r n e m
French	ʁ a - - s i n -			k w a - - - n -	p - o - - - -	ʃ ɛ ʁ - - -
Spanish			r - i s t ɾ a		p j e - l - -	k a ɾ n e -
Italian		f u n - e -	r - ɛ s t - a	k - o t e n n a	p - ɛ l l e -	k a r n e -
PortugueseBrazil			ʁ - ɛ ʃ ʧ - i	k - o ʤ i - - ɐ	p - ɛ - l i -	k a ʁ n i -
Portuguese			ʁ - e ʃ t - ɨ	k - o d i - - ɐ	p - e - l ɨ -	k a ɾ n ɨ -
Catalan				k - o - - n n ə	p - e - ʎ - -	k a r n - -
Walloon	r ɛ - - s ɛ n -			k - oj i - n - -	p ɛ - - - - - -	ʧ ɑ r - - -
Friulian			r j e s t - e		p j ɛ - l - -	c aː r - - -
Romanian	r ə d ə ʧ i n ə	f u n j e -			p j e - l e -	k a ɾ n e -

Character Assignments

Each segment gets a different number

1 i 2 n 3 f 4 a 5 t 6 e 7 m 8 ɑ̃ 9 ĩ 10 ɐ̃ 11 ʧ 12 ɨ 13 ɛ 14 ɲ 15 o 16 ẽj̃ 17 p 18 s 19 k 20 θ

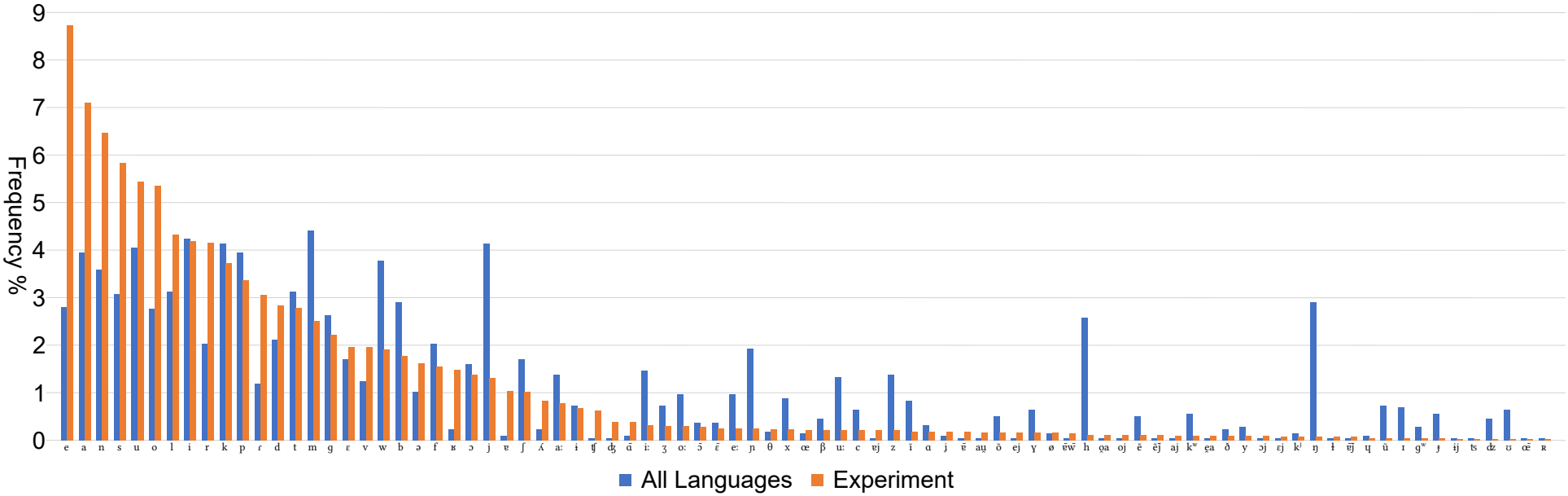
21 ʃ 22 ej 23 ɐj 24 w 25 ɔ̃ 26 h 27 õ 28 b 29 v 30 au̯ 31 l 32 u 33 z 34 ʎ 35 j 36 r 37 x 38 ɾ

39 ɐ 40 ə 41 ɛ̃ 42 ɐ̃w̃ 43 c 44 ɨj 45 eː 46 d 47 ɔ 48 oː 49 œ 50 ɔj 51 o̯a 52 ʤ 53 aː 54 iː 55 ʁ

56 uː 57 ʤ 58 oj 59 ɑ 60 gʷ 61 g 62 ŋ 63 ø 64 ɟ 65 β 66 ʀ 67 ɛj 68 kʲ 69 ɪ 70 ẽ 71 ʒ 72 ɟ 73 ɣ

74 ð 75 e̯a 76 kʷ 77 aj 78 ɥ 79 y 80 ʦ 81 ɬ 82 ʊ 83 œ̃ 84 ũ 85 ɐ̃j̃

Prior Segment Frequencies



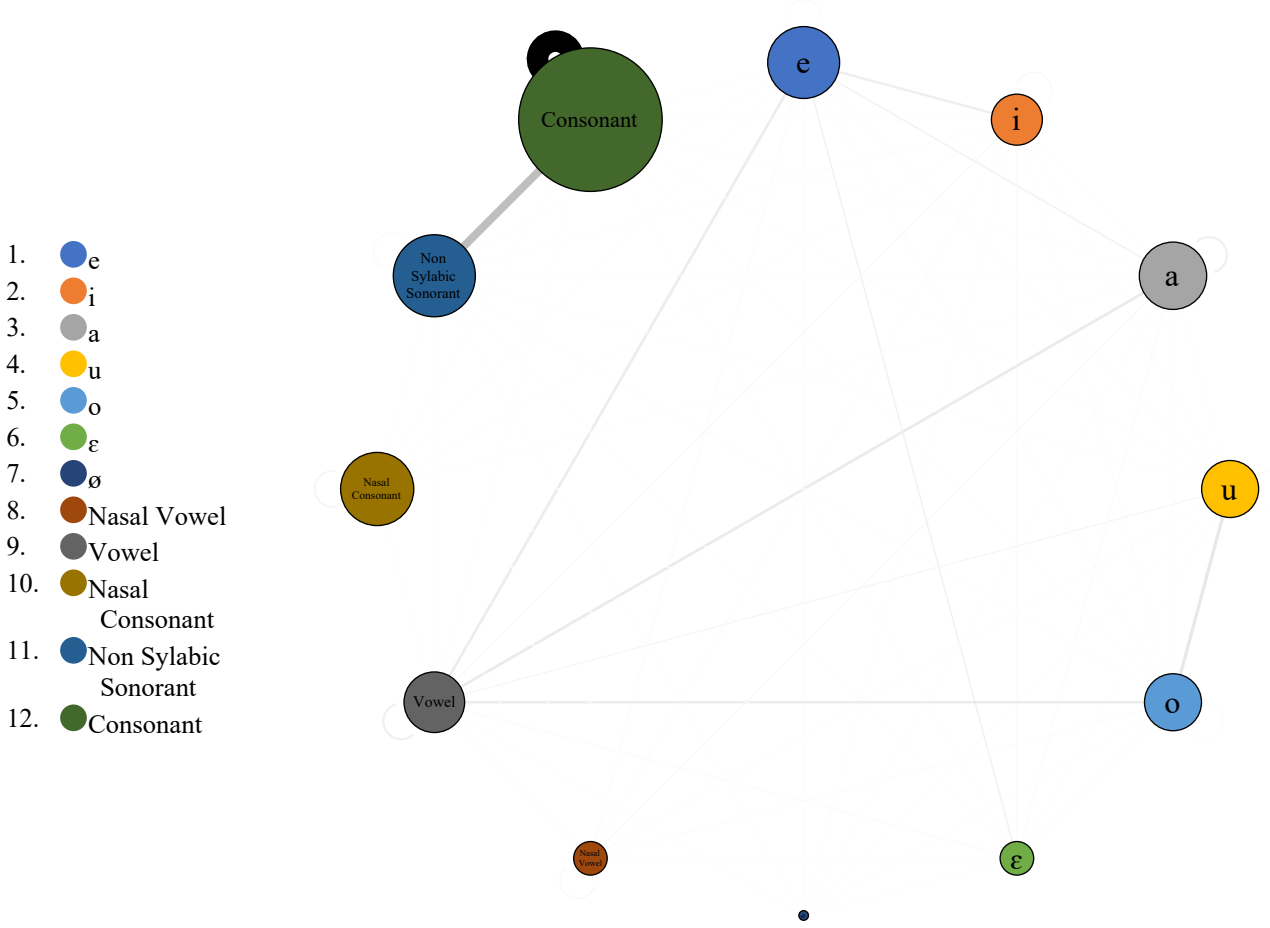
Partition Assignments

Model: “Linguistically Informed”

1)	e	e eː
2)	i	i iː
3)	a	a aː
4)	u	u uː
5)	o	o oː
6)	ɛ	ɛ
7)	ø	ø
8)	Nasal Vowel	ɑ̃ ĩ ɐ̃ ɛ̃j ɔ̃ õ ɛ̃ ɐ̃w̃ ẽ œ̃ ũ ɐ̃j
9)	Vowel	ɨ ej ɐj au̯ ɐ ə ɨj ɔ œ ɔj o̯a oj ɑ ɛj ɪ e̯a aj y ʊ
10)	Nasal Consonant	n m ɲ ŋ
11)	Non Sylabic Sonorant	w l j ɾ
12)	Consonant	f t ʧ p s k θ ʃ h b v z ʎ x ɾ c d ʣ ʁ ʤ gʷ g ʝ β ʁ kʲ ʒ ɟ ɣ ð kʷ ɥ ʦ ɬ

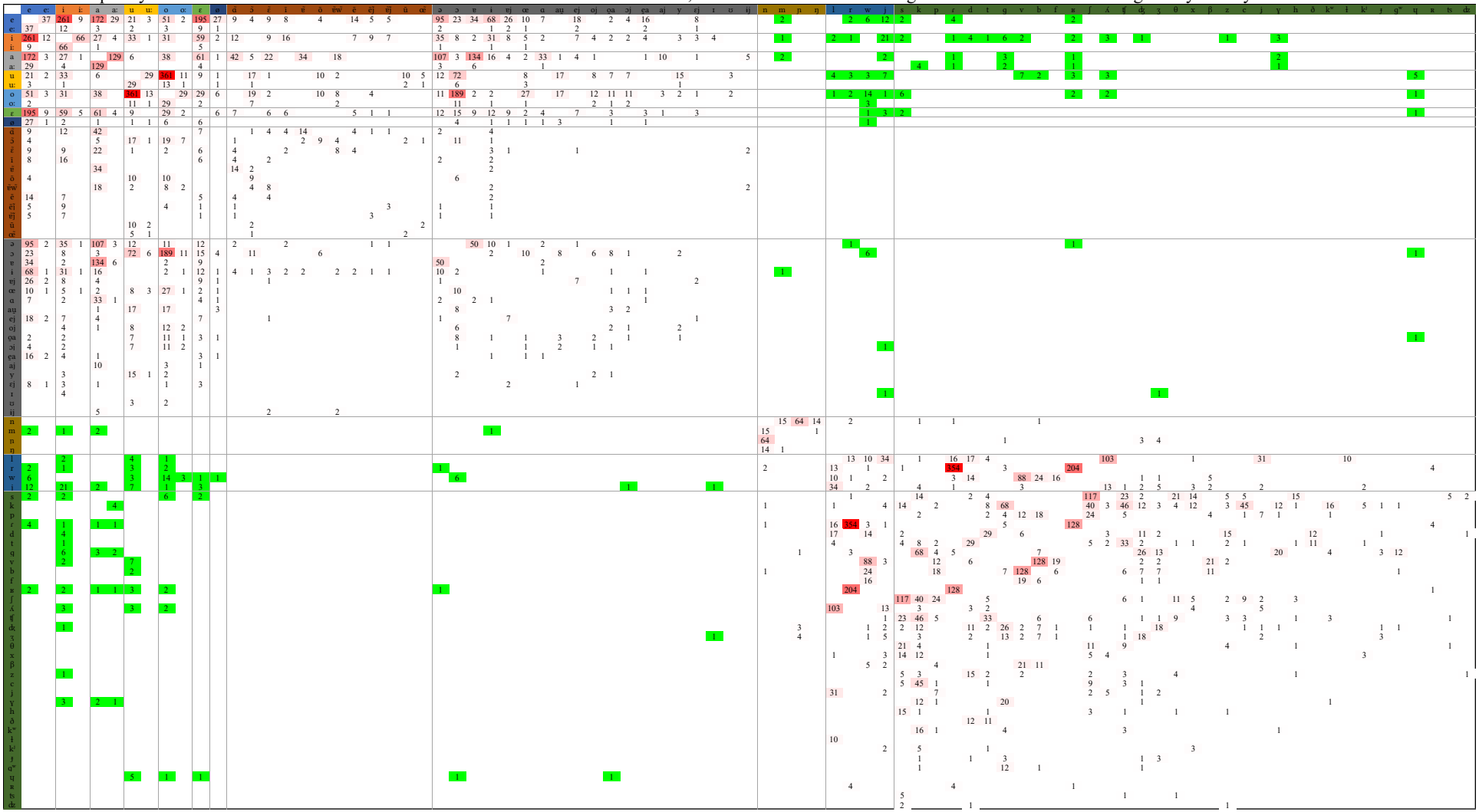
Prior Transition Rates

For the ‘Linguistically Informed’ model, states were grouped into 12 sets. Here, the area of the circles is proportional to the occurance frequencies for each group. The width of the lines is proportional to the rates of transition between each partition.



Segment Transition Rates

This chart shows the number of times that each segment (on the left vertical bar) differs from any other other segment in the same cognate set at the same aligned segment position. The rates of transitions to gaps and from gaps are ignored. The segments are grouped into partitions, represented by the sections divided by lines. Within each partition, the segments are ordered from highest to lowest frequency of occurrence. Transition counts are shown with the number in a red box, where the darker reds are larger numbers. The chart has diagonal symmetry.



The majority rule consensus tree

Questions