Understanding Language Evolution Using an Event-Based Model

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

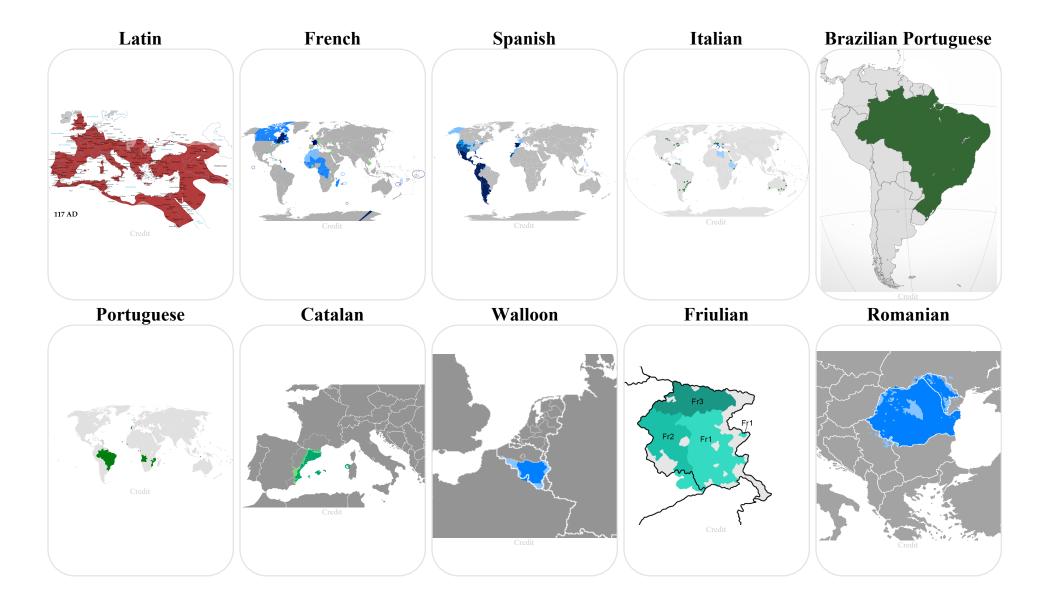
Introduction

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	/mẽ/	1
Spanish	/mano/	1
Italian	/ma:no/	1
Russian	/rʊka/	2
Polish	/rɛŋka/	2

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

Languages



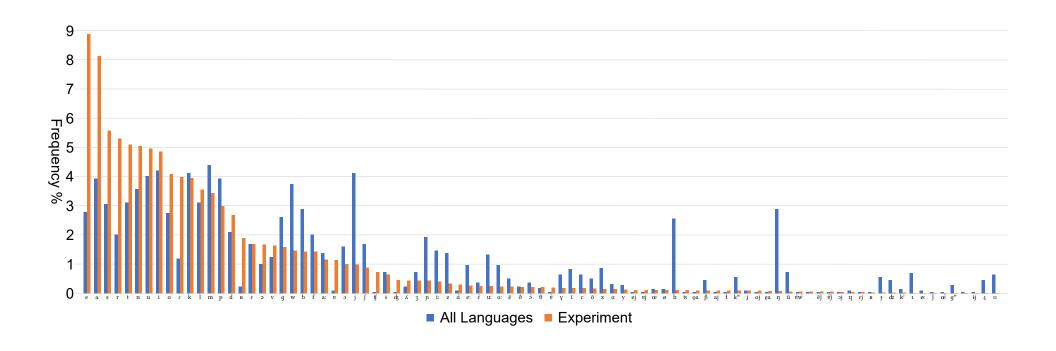
Concept Examples

	Woman-Primary	Woman-Mulier	Woman-Domina	Child-Primary	Child-Ninnus	Wife-Primary
Latin	f e: - m i - n - a	m u l i e r e m	d o m i n a	i n f a n t e m		- s p o n s a
French	f a - m		d a m	$\left \left \widetilde{\mathbf{a}} \right - \left \mathbf{f} \right \left \widetilde{\mathbf{a}} \right - \left - \right - \left - \right - \right $		e - p u - z -
Spanish	- e - m - b r - a	m u x - e r	d o n a	i n f a n t e -	n i n o	e s p o - s a
Italian	f e m m i - n - a	тоббе	d o n - n a	f a n t e -	n i n o	$- \mathbf{s} \mathbf{p} \mathbf{o} - \mathbf{z} \mathbf{a}$
PortugueseBrazil	f e - m e e	т и Л - е в	d o n e	î - f ê - t i -	n e n ẽj	i s p o - z a
Portuguese	f e - m j e	m u Л - e г	d o n e	$ \tilde{\mathbf{i}} - \mathbf{f} \tilde{\mathbf{e}} - \mathbf{t} \mathbf{i} -$	n e n ej	i f p o - z e
Catalan	f e - m - b r - a	m u Л - e	d o n ə		n e n -	ə s p ə - z ə
Walloon	f œ - m		d a m	$ \epsilon $ - $ f $ $ \tilde{a} $ - $ - $ - $ - $		1
Friulian	$f \mid \epsilon \mid - \mid m \mid i \mid - \mid n \mid - \mid e$	m u i: r	d o n e		n i n i	- s p o - s e
Romanian	f e - m e j e	m u - j e r e -	d oam - n ə			1 12 1 1 1 1
	Husband-Primary	Husband-Sponsus	Husband-Socius	Mother-Primary	Father-Primary	Father-Tata
Latin	m a r i: t u s	- s p o n s u s	$s \mid o \mid k \mid i \mid u \mid s$	m a: t r e m	p a t r e m	t a t a
French	такі	e - p u		$ \mathbf{m} \mathbf{\epsilon} - \mathbf{R} - -$	$ \mathbf{b} \mathbf{\epsilon} - \mathbf{R} - $	
Spanish	mariðo-	e s p o - s o -		m a ð r e -	paðre-	t a t a
Italian	m a r i t o -	- s p o - z o -		m a d r e -	padre-	t a t a
PortugueseBrazil		i s p o - z u -		m ej	p aj	t a t a
Portuguese	meriðu-	$\frac{1}{2} \int p o - z u -$		m ej	p aj	t a t a
Catalan	m a r i t	a s p a - s - -		ma-rə-	p a - r ə -	
Walloon				m e - r - -	$ \mathbf{p} \mathbf{\epsilon} - \mathbf{r} - \mathbf{r} $	
Friulian	m a r i: t	- s p o: - s		m a - r i -	p a - r i -	t a t e
Romanian	m ə r i t		s o ts			t a t ə
	Animal-Primary	Animal-Bestia	Fish-Piskem	Fish-Pescion	Bird-Avem	Bird-Aucellus
Latin	a n i m a: 1 i a	b - e: s t i a	p i s k e m		a w e m	- a k - e 1 1 u s
						u u
French	o m a - j -			$p w a s \tilde{\mathfrak{I}}$		
Spanish	a 1 i m a n - a		$p \mid e \mid \theta \mid - \mid - \mid - \mid$		a b e -	
Italian		$\begin{vmatrix} b \end{vmatrix} - \begin{vmatrix} i \end{vmatrix} \int \begin{vmatrix} \int - a \end{vmatrix}$	$p \mid e \mid \int \mid \int \mid e \mid -$			- u t tf ε 1 1 o -
PortugueseBrazil		b - i - f - e	p ej ∫ - i -		a v i -	
Portuguese		$ b $ - $ i $ - $ \int - e $	p ej		a v i -	
Catalan			$ \mathbf{p} \mathbf{e} \mathbf{f} - \mathbf{f} - \mathbf{f} $		a w - -	- u s - e - 1
Walloon		b i ε s - - -		$ p - \epsilon h \tilde{o} $		- u h - ε
Friulian	- n e m a: 1	b - i s - - -	$p \mid \epsilon \mid s \mid - \mid - \mid - \mid$			- u tf j ε - 1 - -
Romanian	- n ə m a - j e		$ \mathbf{p} \mathbf{e} \int \mathbf{t} \mathbf{e} -$			

Character Assignments

Each segment gets a different number

Prior Segment Frequencies



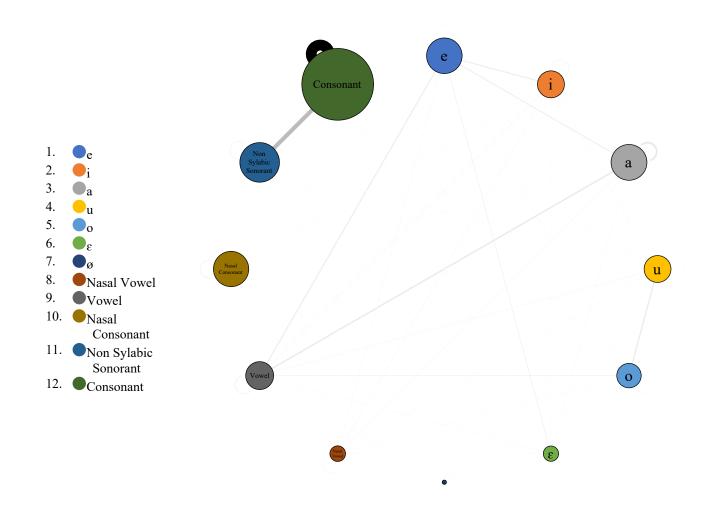
Partition Assignments

Model: "Linguistically Informed"

```
1) e
                                                  e: e
                                                  i i:
                                                   a a:
    u
                                                   u u:
                                                   O O
                                                  3
     Ø
                                                  ø ø:
     Nasal Vowel
                                                  ã î ẽ ẽj ẽj õ õ ẽ ẽw ẽ ũ œ
     Vowel
                                                  υ ω ο ο ο ο e i aj ej ej aŭ ij α y iw oj oj ea εj ι σ
10) Nasal Consonant
                                                  mnnŋĵ
11) Non Sylabic Sonorant
                                                  jlr w
                                                  f \ b \ r \ x \ \land \ b \ d \ t \ f \ s \ p \ z \ \int \check{\eth} \ k \ ts \ \theta \ h \ v \ c \ \beta \ \dot{l} \ u \ g \ dz \ \dot{j} \ dg \ g^w \ R \ k^j \ z \ J \ \gamma \ k^w \ J
12) Consonant
```

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.

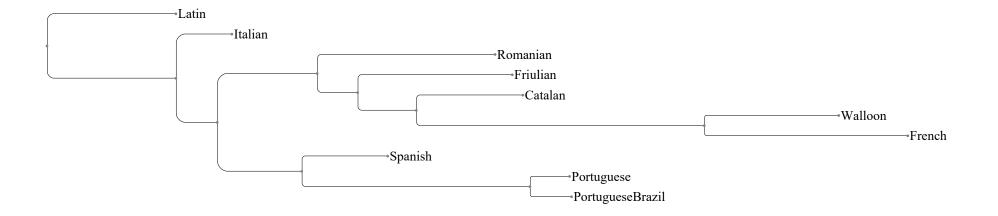


Segment Transition Rates

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 occurrance. Transition counts are shown with the number in a red box, where the darker reds are larger numbers. The chart has diagonal symmetry.

e e i ki a a u u u o o e i a a a a i a e i e i a a o e i a a a a a a a a a a a a a a a a a a	314 79 29 176 11 5 32 36 17 2 9 26 9 13 1 1 1 1 1 2 4 2 5 5 2 5 1	Treas are larger numbers. The chart has diagonal symmetry.
i 693 32 224 83 13 60 2 44 139 6 1 24 56 26 61 2 8 10	126 5 11 73 4 17 12 11 10 2 2 11 8 3 5 4 1 4 9	1 8 33 3 1 1 2 4 2 2 10 1 5 1 3
a 581 8 83 3 554 12 111 138 3 1 88 26 2 7 99 26 8 10 1 1 3 2 2 10 1 3 3 2 2 10 1 3 3 3 3 3 14 14 3 5 3 43 2 32 2 32 8	301 436 34 40 72 7 10 3 4 24 1 5 1 1 5 4 5 13 5 1 1 1 5 4 22 2 2 151 58 9 15 9 17 8 3	
u: 4 2 101 22 1 2 1 1 2 1 0 75 2 44 111 840 22 73 36 13 2 2 2 48 46 8 6 4	10 11 4 26 7 323 6 3 37 26 3 23 21 14 1 2	2 1 2 1 6 4 4 2 2
0 3 3 1 73 2 3 1 7 7 2 3 1 7 7 2 3 1 7 7 2 3 1 7 7 2 3 1 7 7 7 2 1 7 7 7 7 7 7 7 7 7 7 7 7 7 7	26 1 362 0 3 37 20 5 25 21 14 1 2 25 12 20 26 6 3 19 14 3 4 11 12 3 3 6 1 2 1 2 2 1 2 4 1 4	3
2 1 1 2 1 2 1 2 1 3 3 1 2 1 1 8 2 0 1 28 7 1 1 1 3 1 2 56 6 26 1 5 2 19 8 10 12 10 2	1 1 1 4 10 4 10 5 1 6 5 1 2 1 2	
ē 79 26 2 3 15 20 10 3	9 5 4 15 3 4 3 8 1	■
1 17 61 5 2 10 7 12 3	3 4	
û 9 1 8 8 2 6 3 2 4 5 5	2 2 2 1 1 1 1	
0 7 1 10 8 1 4 1 1 5 0 314 11 126 5 301 5 22 26 25 10 5 9 4 4 3 1 1 7 79 2 5 436 13 2 7 12 4 3	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
29 11 34 151 10 323 28 20 6 1 15 10 4 1	117	•
a 11 2 4 7 72 5 58 11 3 6 2 1	6 6 1 1 1 1 1 1 1 1 2 1 2 1 1 1 1 1 2 1 2	<u> </u>
ei 36 5 11 10 1	4 1 1 10 10 1	_
coa 2 2 4 15 26 5 4 2 aj 9 2 1 24 3 11 11 4 4 oj 111 1 9 23 8 14 4 4	2 1 1 3 2 3 1 2 1 7 2 2 2 2	<u> </u>
au 1 17 21 1 4 4 4 4 1 4 4 1 4 4 1 4 4 1 4 4 1 4 4 1 4 4 1	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	_
i 13 1 5 1 1 3		-
1	41 185 3	5 9 9 2 7 1 4 1
2 9 2	41 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
	9 9 2	10 54 6 1 1 12 11 4 130 3 12 2 1 1 1 14 10 54 95 79 4 4 2 1 21 4 6 156 3 48 35
w 6 31 24 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	3 2	6 95 4 8 10 18 183 48 22 2 1 2 1 7 6 1 1 79 4 1 4 3 7 21 17 12 02 2 2 2 5 1 7 2
	7	1 12 4 40 41 3 1 1 1 24 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
	4 1	5 11 31 3 37 38 5 2 1 11 7 2 4 21 18 5 144 6 4 4 2 43 3 13 22 41 1 1 1 1 2
y 2 7 7 9 5 10 3 2	4 1	730 4 3 430 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
	1	48 2 37 4 310 10 7 6 7 7 20 1
1 1 ds 1 5 5 2 2	3	1 117 136 118 5 2 6 7 13 1 1 23 30 8 11 4 10 1 3 1 1 2 1 1 3 16 1 2 43 4 2 45 7 1 1 1 93 4 3 3 1 9 1 5 1 1 1
3 2 1	4	1 22 1 5 13 2 40 7 1 1 93 4 5 13 8 3 2 2 2 2 62 2 9 22 2 2 23 4 4 10 6 1 1
0 0 0 0 1 1 1 1 1 1		66 41 3 3 10 4 3 2 1 2 1 3 4 2 61 3 5
c x		5 10 100 1 1 14 8 1 3 5 14 1 13 1 1 5 9 12 13 19 42 3 1 11 4 1
B B B B B B B B B B B B B B B B B B B		1 1 26 1 2 9 4 1 6 3
		1 48 6 1 35 7 7 1 2 5 4 8 1 8 56 2 1 3 10 2 2
14 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	• •	14 8 4 1 12 1 4 1 1 3
		2 8 1 5 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
		1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

The majority rule consensus tree



Questions

