Understanding Language Evolution Using an Event-Based Model

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

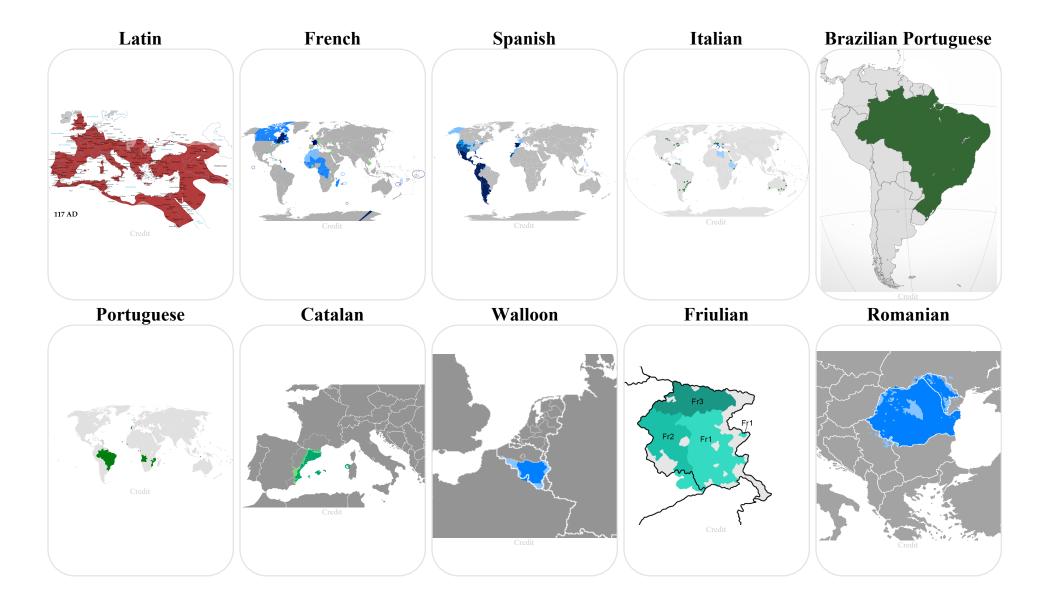
Introduction

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



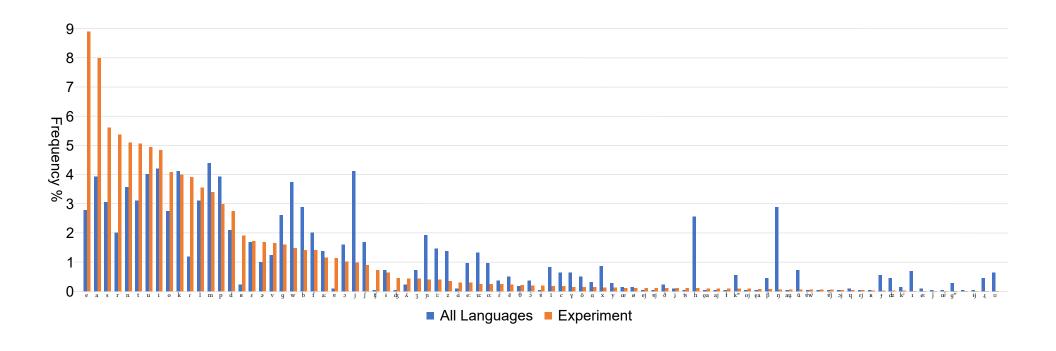
Concept Examples

	Woman-Primary	Woman-Mulier	Woman-Domina	Child-Primary	Child-Ninnus	Wife-Primary
Latin	f e: - m i - n - a	m u 1 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	$- \mathbf{s} ^{2} \mathbf{p} \mathbf{o} - \mathbf{z} \mathbf{a}$
PortugueseBrazil	f e - m e e	т и А - е в	d o n e	î - f ê - t i -	n e n ẽj̃	i s p o - z a
Portuguese	f e - m j e	m u K - e r	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 K - e	d o n ə		n e n -	ə s p э - z ə
Walloon	f œ - m		d a m	$\mid \epsilon \mid$ - $\mid \mathbf{f} \mid \tilde{\mathbf{\alpha}} \mid$ - \mid - \mid -		
Friulian	$f \mid \epsilon \mid - \mid m \mid i \mid - \mid n \mid - \mid e$	m u i: r	d o n e		$ \mathbf{n} \mathbf{i} \mathbf{n} \mathbf{i} $	- s p o - s e
Romanian	f e - m e - - j e	m u - j e r e -	d oam - n ə			
	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 o k i u s	m a: t r e m	p a t r e m	t a t a
French	m a k i	e - p u		m ε - R	b ε - R	
Spanish	marido-	e s p o - s o -		madre-	padre-	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	m e r i ð u -	i ∫ p o - z u -		m ej	p aj	t a t a
Catalan	m a r i t	ə s p ə - s		ma-rə-	pa-rə-	
Walloon				m e - r	p ε - 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: l 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
French	o ma - j -			$p \mathbf{w} \mathbf{a} \mathbf{s} \tilde{\mathfrak{I}}$		ŭ .
Spanish	0 - - III a - J -		p e 0 - - -	p w a s s	a b e -	w a z - o
Italian		$ \mathbf{b} - \mathbf{i} \int \mathbf{f} - \mathbf{a} $	$p \mid e \mid b \mid = = = = = = = = = = = = = = = = =$		a b e -	- u t t ε 1 1 o -
		b - i - f - e			a v i -	$ - u t y \epsilon 1 1 0 -$
PortugueseBrazil Portuguese		$\begin{vmatrix} \mathbf{b} & - & \mathbf{i} & - & \mathbf{j} & - & \mathbf{e} \\ \mathbf{b} & - & \mathbf{i} & - & \mathbf{j} & - & \mathbf{e} \end{vmatrix}$	p ej ∫ - i - p ej ∫ - i -		a v 1 - a v i -	
Catalan		n - 1 - 1 - 6	1 3 0			- u s - e - K - -
Walloon		b i ε s - - -	p e f - - -	$ p - \epsilon h \tilde{o} $	a w - -	- u s - e - Λ - u h - ε
Friulian	- n e m a: 1		n 0 0	$ \mathbf{p} - \mathbf{\epsilon} \mathbf{n} 0 $		
		b - i s - - -	$p \mid \epsilon \mid s \mid - \mid - \mid - \mid$			$-\left u\right tf\left j\right \epsilon\left -\left 1\right -\left -\right $
Romanian	- n ə m a - j e		p e t e -			

Character Assignments

Each segment gets a different number

Prior Segment Frequencies



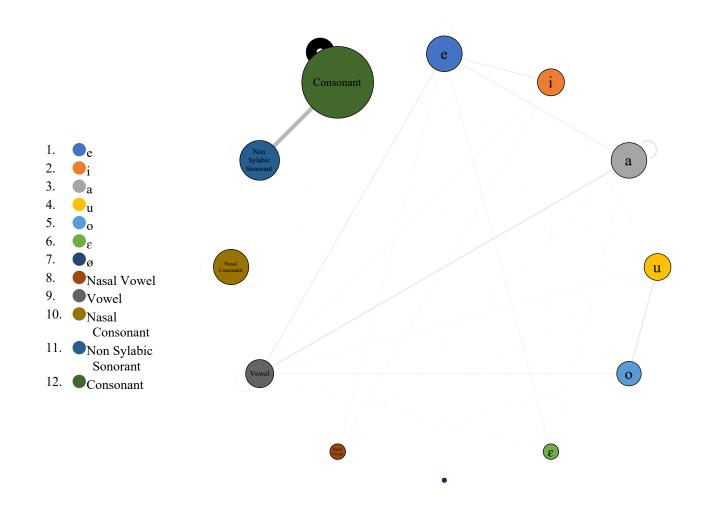
Partition Assignments

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 \ b \ r \ x \land \ b \ d \ t \ f \ b \ 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.

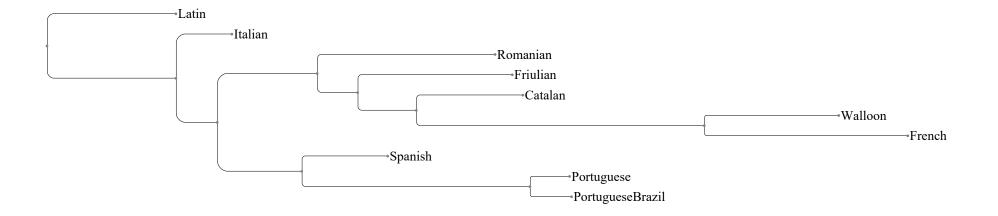


Segment Transition Rates

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.

	shown with the number in a red box, where the darke	er reds are larger numbers. The chart has diagonal symmetry.
e e e i i a a a 1 1 1 0 0 0 8 0 1 3 1 5 1 1 0 1 1 1 1 0 0 1 3 50 1 3 50 1 3 1 7 9 8 1 7 4 7 9 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	312 79 29 174 11 5 32 17 36 2 9 26 8 13 2 11 11 2 4 2 5 2 5 5 5 1	2 6 17 4 2 5 3 3
i 686 33 223 81 13 58 2 43 143 6 1 24 56 26 61 2 10 8	125 5 11 73 4 17 12 10 11 2 2 11 8 3 5 4 1 4 9	1 8 38 3 1 2 1 4 2 2 9 1 5 1 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
a 563 9 81 3 542 11 107 138 3 1 86 26 2 7 99 26 8 a: 128 13 542 1 1 10 10 1 3 2 2	298 433 34 40 72 7 3 10 4 24 1 5 1 1 4 5 2 5 13 5 1 1 1	2 4 1 1 2
n: 4 2 99 22 1 2 1 1 2 1	22 2 152 57 9 15 9 17 8 3 10 11 4	3 4 24 9 4 2 2 3 7 2 3 14 2
o 75 3 43 107 834 22 73 37 13 2 2 46 44 6 8 4 oc. 3 1 73 2 3 1 2 3 1 2 2 4 4 6 8 4 2 x 504 26 143 13 13 10 14 37 2 10 2 11 19 15 10 1 3	26 7 323 6 3 37 26 3 23 21 14 1 2 2 8 1 2 20 26 6 3 19 3 14 4 11 12 3 3	2 1 24 1 6 4 4 2 2 1 3 1 3 1 3 1 3 1 3 1 3 1 3 1 3 1 3
e 504 26 143 13 138 10 14 37 2 10 2 11 19 15 10 1 3	25 12 20 26 6 3 19 3 14 4 11 12 3 6 1 2 1 2 1 2 2 4 1 4	
1 33 5 6 6 26 1 5 2 19 8 8 10 12 2 10 1	10 4 10 5 1 6 5 1 1 2 2	
5 8 26 2 3 15 20 10 3 19 2 4 1	9 5 4 15 3	
10 28 2 7 12 3	4 3 8 1 1	
0 4 26 34 1 19 39 2 44 1 19 39 2 4 4 3 19 30 2 4 4 4 3 19 30 2 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	10 4	
ivi 2 26 2 2 8 2 1 10 4 ij 7 1 10 8 8 2 1 1 1 5 ij 9 1 8 8 2 3 1 1 5	1 1 1	
8 1 1 4	1 1 1 36 6 2 3 4 2 1 1	
312 11 125 5 298 5 22 26 25 10 5 9 4 4 3 1 1 1 25 29 29 11 34 152 10 323 28 20 6 1 15 10 4 1	117 1 6 1	
1 174 4 73 3 40 6 1 26 1 1 10 6 5 3 8 4 2 1 1 6 2 1 1 1 2 4 72 5	36 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
	3 1 1 2 2	· · · ·
ac 17 2 10 1 3 9 4 37 5 3 1 3 6 5 11 10 12 14 2 1 14 2 1 1 9 2 2 4 15 26 5 4 2 2 2 9 2 1 24 3 11	3 15 3 1 1 1 10 22 1 1 3 2 3 1	<u> </u>
aj 9 2 1 24 3 11 2 2 1 1	2 1 1 2 2 1 1 2	
ea 26 5 8 5 12 1	1 1 1 1	
ei 13 1 5 1 1 3	1 2 1 1 2 2 1 1 1 2 1 1	<u> </u>
1		
5 3 2 2	41 177 2	9 9 6 2 1 7 4 1
m 2 9 2	41 5	2 1 1 3 4
	177 5 29 2	10
7 2 1 8 2 1 1 <u>3 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 </u>	9 6	0 54 6 1 1 12 10 10 14 132 3 12 2 1 1 1 14 15 10 54 95 74 4 1 42 21 4 6 157 3 41 48 6 95 4 10 8 18 184 22 49 2 1 2 1 6 5 1
w 6 24 24 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	8 2 1 1	6 95 4 10 8 18 184 22 49 2 1 2 1 6 5 1 1 74 4 4 1 3 7 1 11 18 22 2 2 5 7 1 2 2
		$ \begin{array}{cccccccccccccccccccccccccccccccccccc$
	7	1 10 4 41 29 11 134 77 118 12 3 5 9 11 100 26 13 3 2 56 12 5 1 1 42 8 1 3 416 5 1 1 7 12 7 1 18 12 7 1 18 12 7 1 18 12 7 1 18 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
d 4 k 3 2 1 1 3 2	4 1	4 21 18 5 171 6 4 4 2 41 3 11 22 1 35 1 1 2
y 3 2 7 7 3 2 7	4	3 7 1 3 134 5 4 4 2 10 46 41 1 51 1 1 4 3 12
b 2 1	1	
	3	49 2 40 4 317 10 7 6 7 7 13 1 3 1 1 2 15 14 77 38 2 13 1 1 2 15 14 5 1 2 9 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 3 2 2	4	157 18 2 3 3 1 12 6 1 1 1 2 1 1 7 1 93 4 3 12 8 3
		2 2 2 62 2 9 22 2 2 23 4 4 10 1 6 1
c y 3 2 1		5 10 100 1 1 1 4 8 1 2 1 26 1 51 1 3 1
ă		3 5 14 1 13 1 5 7 12 12 13 3 1 5 7 12 12 13 13 1 5 7 12 12 13 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
		41 7 1 26 1 2 1 1 1 4 1 1 1 1 4 1 1 1 1 1 1 1 1 1
i l		1 1 26 1 2 9 4 1 6 3 1 48 6 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
β 14 2 1 1		5 2 7 23 13
	_ _ _	14 8 4 1 12 1 3 1 1 3
dz k ⁱ		8 1 2 3 1 1 1 2
9		1 12 1 1

The majority rule consensus tree



Questions

