# Understanding Language Evolution Using an Event-Based Model

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

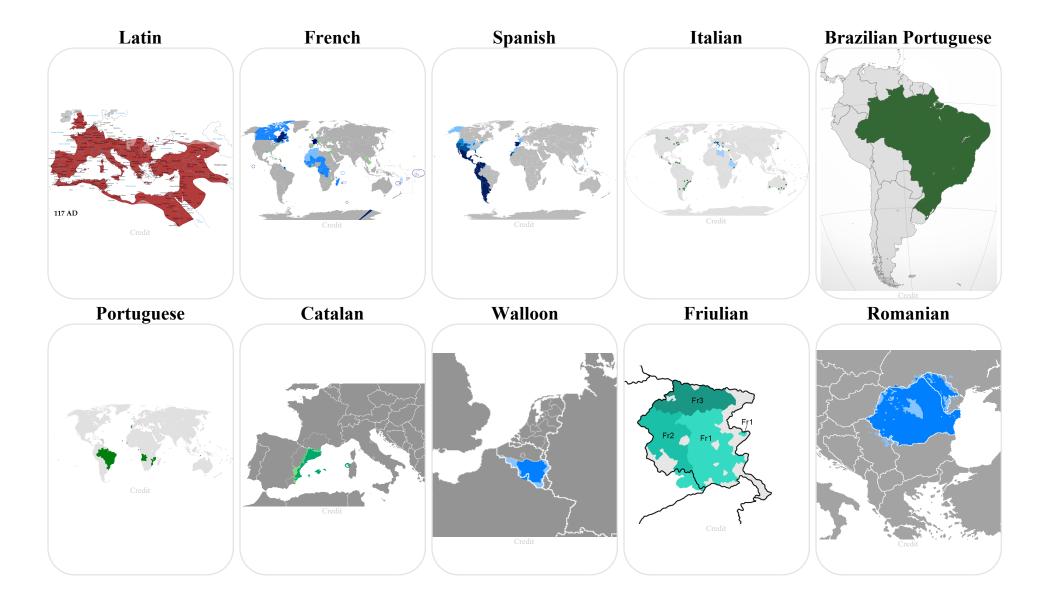
#### Introduction

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



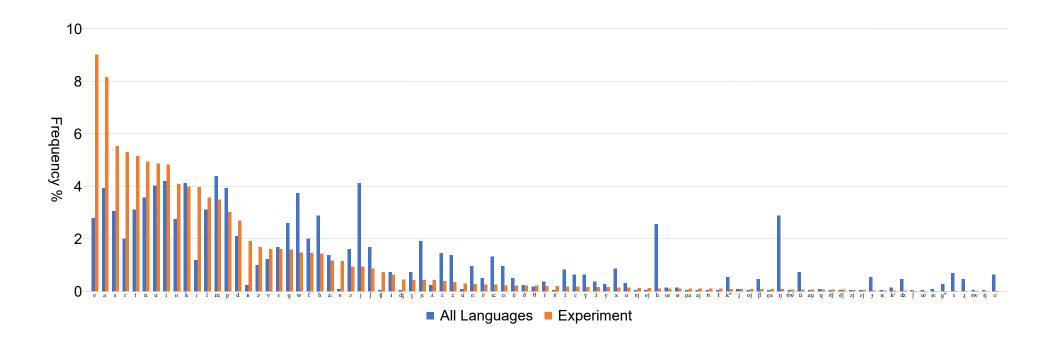
# Concept Examples

	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	$ \tilde{\mathbf{a}} $ - $ \mathbf{f} $ $ \tilde{\mathbf{a}} $ - $ \mathbf{f} $ -		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	- s p s - z a
PortugueseBrazil	f e - m e e	т и л - е в	d o n e	î - f ê - tf i -	n e n ẽĩ	i s p o - z a
Portuguese	f e - m j e	m u k - e r	d o n e	î - f ê - t i -	n e n ej	i f p o - z e
Catalan	f e - m - b r - a	m u \lambda - e	d o n o		n e n -	$\begin{vmatrix} \mathbf{a} & \mathbf{s} & \mathbf{p} \\ \mathbf{s} & \mathbf{p} \end{vmatrix} \mathbf{s} - \mathbf{z} \begin{vmatrix} \mathbf{a} \\ \mathbf{s} \end{vmatrix} \mathbf{e}$
Walloon	f œ - m		d a m	$\mid \epsilon \mid$ - $\mid f \mid \tilde{\alpha} \mid$ - $\mid$ - $\mid$ - $\mid$ -		
Friulian	$f \mid \epsilon \mid - \mid m \mid i \mid - \mid n \mid - \mid e$	m u - - i: r - -	d o n e		n i n i	-   s   p   o   -   s   e
Romanian	f e - m e - j e	mu - 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	t a t a
French	тав і	e   -   p   u   -   -   -   -		те-в	р е - в -	
Spanish	mariðo-	e s p o - s o -		m a ð r e -	раðге	t a t a
Italian	m a r i t o -	- s p o - z o -		m a d r e -	p a d r e	t a t a
PortugueseBrazil	m a r i d u -	i s p o - z u -		m ej	p aj -   -   -	t a t a
Portuguese	m e r i ð u -	$ \mathbf{i}  \int  \mathbf{p}  \mathbf{o} -  \mathbf{z}  \mathbf{u} -$		m ej	p aj -   -   -	t a t a
Catalan	marit-	ə s p ə - s		m a - r ə -	ра - г э	
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-Aucellum
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 m a - j -			p w a s õ		w   a   z   -   o   -   -   -   -
Spanish	a   1   i   m   a   n   -   a		p e θ		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 t ε 1 1 o -
PortugueseBrazil		b   -   i   -   f   -   e	p  ej ∫   -   i   -		a   v   i   -	
Portuguese		b   -   i   -   f   -   e	p ej ʃ - i -		a v i -	
Catalan			p e		a   w   -   -	-   u   s   -   e   -   \lambda   -   -
Walloon		b   i   ε   s   -   -   -		$ p  -  \epsilon  h  \tilde{o} $		-   u   h   -   ε   -   -   -   -
Friulian	- n e m a: 1	b   -   i   s   -   -   -	$p \mid \epsilon \mid s \mid - \mid - \mid - \mid$			$-\left \mathbf{u}\left \mathbf{tf}\right \mathbf{j}\left \mathbf{\epsilon}\right -\left 1\right -\left \mathbf{r}\right $
Romanian	- n ə m a - j e		p e f t e -			

#### Character Assignments

Each segment gets a different number

# Prior Segment Frequencies



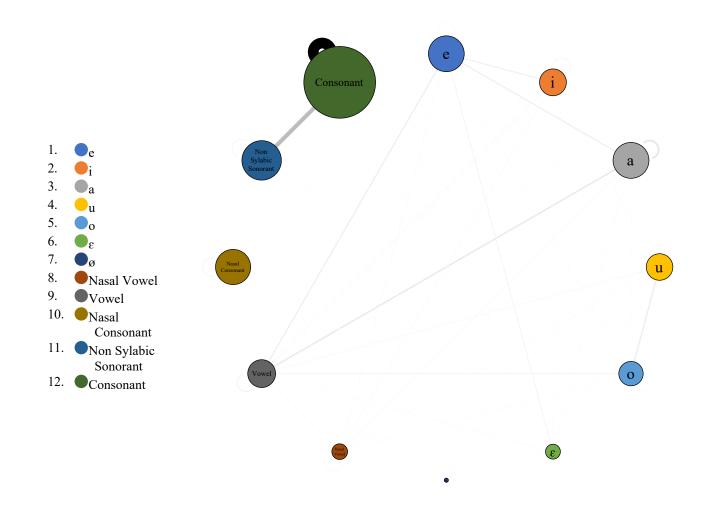
### Partition Assignments

#### Model: "Linguistically Informed"

```
1) e
                                                                                                                                                                                                                                                                                                                                                                                                                                    e: e
                                                                                                                                                                                                                                                                                                                                                                                                                                  i i:
                                                                                                                                                                                                                                                                                                                                                                                                                                     a a:
                                        u
                                                                                                                                                                                                                                                                                                                                                                                                                                     u u:
                                                                                                                                                                                                                                                                                                                                                                                                                                     o o:
                                                                                                                                                                                                                                                                                                                                                                                                                                    3
                                                 Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                    ø ø:
                                                                                                                                                                                                                                                                                                                                                                                                                                  \tilde{a} \tilde{i} \tilde{e} \tilde{e} \tilde{j} \tilde{e} \tilde{i} \tilde{e} 
                                               Nasal Vowel
                                            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
12) Consonant
                                                                                                                                                                                                                                                                                                                                                                                                                                 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
```

#### **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.

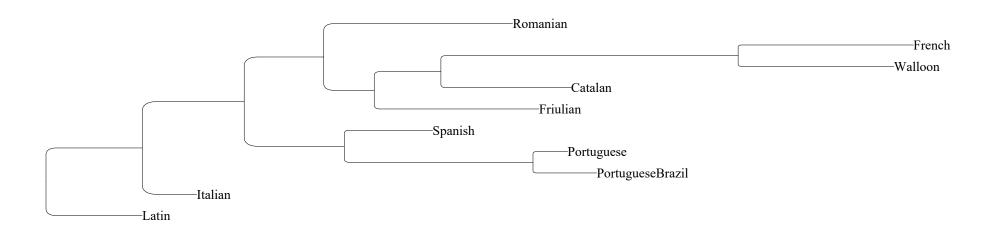


## Segment Transition Rates

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: e 100	i i	a a:	u u: 0 44 6 7 1 1	0:	e 0 0	i a a	õi	p.	i 5	êw û	<b>ēj ēj</b> 11 7 1 1	œ ə	В	28 177 3	y a	ej ej	œ o	a aj	oj ea 26	au o	j εj 13	Ιŭ	iw ij	n m	ր դ	j r	- 1	w j	s t 4 4	k 1	р	д к	v g	f b	_ [ _ [	ds 3	Á Z	ðθ	9 с у	/ x 1	ı ts <del>l</del>	j k'	″βι	Į R	j k <sup>j</sup>	dz g*	[
i 696 28 i: 32 2	219	9 78 13		2 1	36 6 1 13 1	24 3	1 4	7 6	51 5		8 10	12	0 5	10 68	19 9	12 11	10 2	1			5	4	1	4 9			8	36 1	4 1	2		4	2 7			1	5 1										
a 617 12 a: 143 1		565 565	11 11	14 1	29 3 1 11	88 7	1 1	3	2 33	2	8	7	1.4	34 40	66	8 11 1 1	2 4	26	1 5	1	1	3	4 5 1	2		-		22 7	1	4			2	2										-			_
u: 6 1	2	114	102 8. 102 2	2 1	111 12 14 1 2 138 13 2 2 3 1 9 2 9 1	1 2	61 2		2 33 1 37	2	4	1 1	3 2 1	10 11 6	14		3 2	7 3	23	16 1	4 1	2					1	28 1	6 4	4			/		_		2										_
ο: 3 ε 490 24	136 1	3 129 11	35 1 7 12 1 3	7 8 2	2 3 1	11 1			6	2	1 1	2:	5 14	26 1 21 25	3 6	20 15	5 5	9	8 12	1 4	3		1					3 4	5																		
	1	1	14 2 1	3 3	9 1			20						6 1	1 2	2 2	1 2	!	4 1	4																											
ē 83 2	34	88	3 1 4	1	16	20	1	28	5 17	2	1 1	9 2	) 4	7		1 1																															
δ 5 ε 24 3	47 5	17 101 3	5 2	2	18	8 1: 28 7 5			0 2	6 2		5		1 4		4							2	2																							
i 17 5 3	61 5	6	2 33 1 3	7 6	10	7 5	17	2		4 2		1 2		4 11 2										1			•	1																			
ēw ŭ ēj 11 1	2	26 2 8 2	2 8	3 2	2	1 2	2		2		5	4		4									2 2																								
ēj 7 1	10		8 1		1	i			1	4	5	i		i i																																	
ə 313 10 e 79 2	120 4	312 7 438 14	18 1 2	4	25 14		2 5				1 1	12	0	1 35	5	2 4		2 1			1 1						_																				
28 i 177 y 5	10	34 40	145 10 3 66 14	11 26	21 6 25 1 1	10 7	14 1	8	4 2	2 4	1 1	1 1	5	2 2 3 1	3 1 1	1 1	13 2	1 1	7 1 2	8 2	2 1			1				7															_				
a 13 2 ej 31 5	9 1	66 5	00 14 :	,	6 2		4					6	5	1		1 10	1	I	1		2						•	2															_	_			
ej 35 5 œ 16 2	12 11 10	8 1 11 1 2 4	8 3 3	3 5	15 2 1 3 1	i						4		1	1	10	3	;	1		1																										
oa 2 aj 11	10 1 2 2 11	26 1		7 5 3 8	4 2 9 4 4							2	1	21 1 1 7	2		3		2	3 1																											
ea 26 5	8	5			12 1 4									1 8	1		1	1		-	,							2																			
οj 9 εj 13 1	4 5	1	7 1	6 1 4 4	3							1		2		2 1	1	ĺ	2	2	1						·	Ĩ																			
I U	4	١	3 2	2						2																		1								<u></u>											
ij	4	4 1 5			1		- 2		_	2														41	185 35	9	9	2		1 7	,	4		1													
m 2	9	2												1										41 185 1 35 2	1 2		2					1	4 1			3 4											
r 1	8	1	3																					9	2	10 55	10 55 5 6 97 1 68	6 1 97 68	1 12	4 4	5 3	4 <b>735</b> 21 4	6			12	153			4	1 1	8 35		14			-
w j 13	36		32 2	8 3	4				1					7	2					2		1		2		6	97 68	3	20	10 8	3 I	19	181 7	19 49	216 11	2 1	2 2 18 2 62 2 2 3 12 9	- 49	0 0	2 5	1 6	5 1 7 8	5 3	3	2	8	_
t 4 k 2	1 2	1 4	2 2		•																			1		12	2 4	10 4	38 39 30	30	5 1	43 3	3	2	14 13 87 11	8 16 8 12 5	2 2	66 6	6 10 1 1 100 3	1 1 5 13 2	1 41	56			1 12 5	1	
r 3		-	1																					7		12 6	45					435	33 3 6 4	37	38 5				1 2	2	1	1 72	11	8			1
d в 2	4		7																					4	1	73	21 5 4 6		5 143	43	33		6 4	4	2	43 13 4 2 2	3 3 22		1	1		1	25	4	1	2	
g 5	7	2																							1			7	7 3	134	3	4	2 27 317 10	10	22 7	42 40	0		4 5	5 1		1 3			4	12	
Ь			2																					1				49	2 315 14	87	37 38	4	317 10	7 22	6 14	7 7	2	15	5 11	5 9	) 1	2	22		1	1	
虚 1	1																					1			3 4	12	2		114 138 5 16		5	2 43 4 13	2 42 2 40	7 6	1 1	1 1	2 23 3 4 4	3	2 3	3 9 °	1 10 1	5 8	D	1	1 1 3	3	
A Z	5			2																	-	•			+	,	153	1 22 18	5 16 1 2 62 2	12		3	2 40	1 /		4 4			0	15	1	1 4			3	1	
ŏ e																										-			66 68 6	11	4	41 1			15 30	3	10				3 3	. 2	!			1	
c Y	1	-																									,	2		35	1 2		55		11 8	3 5	3 15					2	!		3		
h ts																										1		1	26 1 17 41	13 2 3		1	1		9 4	9 13	5 15	3	3						3	1	
± j																										1	48 35	6 1 7		1	7		1		2	5 8	1 4 1										
β													_				_											5	8	56	2 11		25	1 22	10			2	2 2	2						1	
ų R			1/ 1	•									•	1												14	4	3		12	3	1 4	4		1 1	1 2											
k <sup>j</sup> dz																												2	8 1	5		2			3			1	1	3	1	1					
g" -L																														1	ı		12	1		1											

# The majority rule consensus tree



# Questions

