Understanding Language Evolution Using an Event-Based Model

David Goldstein

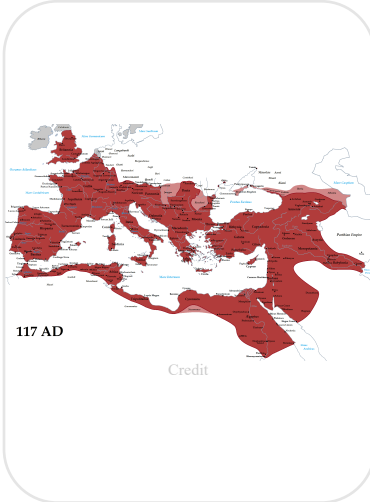Department of Linguistics
dgoldstein@humnet.ucla.edu

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	/rʊka/	2
Polish	/rɛŋka/	2

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

Latin
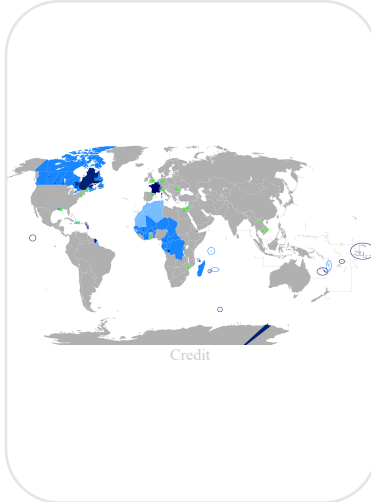
French

Spanish
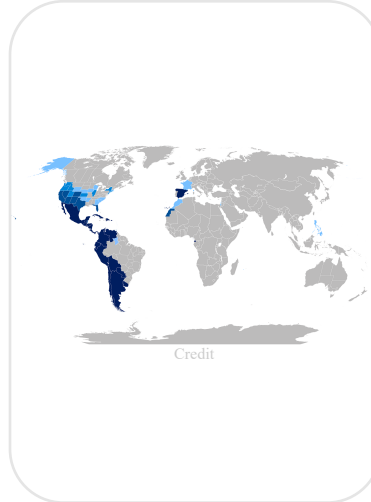
Italian

Brazilian Portuguese

117 AD

Credit

Credit

Credit

Credit

Credit
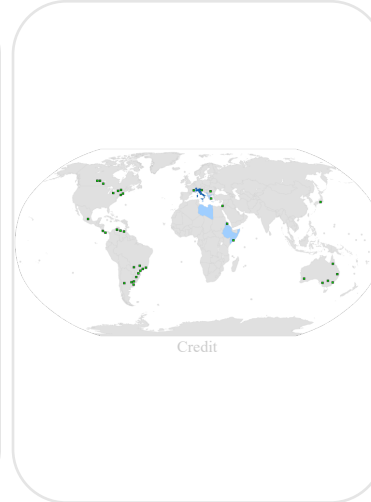
Portuguese

Catalan

Walloon

Friulian

Romanian

Credit

Credit

Credit

Fr3

Fr1

Fr2

Fr1

Credit

Credit
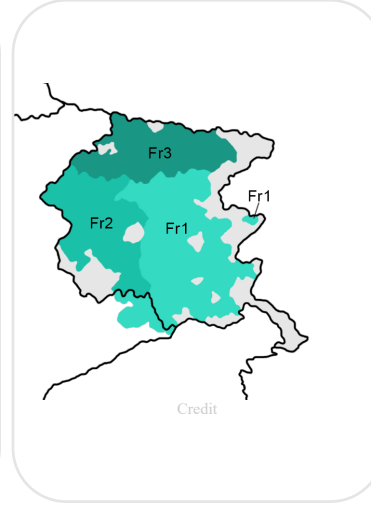
	Woman-Primary	Woman-Mulier	Woman-Domina	Child-Primary	Child-Ninnus	Wife-Primary
Latin	f eː m i n a	m u l i e r e m	d o m i n a	i n f a n t e m		s p o n s a
French	f a m		d a m	ã f ɑ̃		e p u z
Spanish	e m b ɾ a	m u x e ɾ	d o ɲ a	i n f a n t e	n i ɲ o	e s p o s a
Italian	f e m m i n a	m o ʎ ʎ e	d ɔ n n a	f a n t e	n i n o	s p ɔ z a
PortugueseBrazil	f e m e ɐ	m u ʎ e ʁ	d o n ɐ	ĩ f ě ʧ i	n e n ẽj	i s p o z ɐ
Portuguese	f e m j ɐ	m u ʎ e ɾ	d o n ɐ	ĩ f ě t ɨ	n e n ẽj	ɨ ʃ p o z ɐ
Catalan	f e m b ɾ a	m u ʎ e	d ɔ n ə		n e n	ə s p ɔ z ə
Walloon	f œ m		d a m	ɛ f ɑ̃		
Friulian	f ɛ m i n e	m u iː r	d ɔ 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 o̯ a m n ə			

	Husband-Primary	Husband-Sponsus	Husband-Socius	Mother-Primary	Father-Primary	Father-Tata
Latin	m a r iː t u s	s p o n s u s	s o k i u s	m aː t r e m	p a t r e m	t a t a
French	m a ʁ i	e p u		m ɛ ʁ	p ɛ ʁ	
Spanish	m a ɾ i d o	e s p o s o		m a d ɾ e	p a d ɾ e	t a t a
Italian	m a r i t o	s p ɔ z o		m a d r e	p a d r e	t a t a
PortugueseBrazil	m a ɾ i d u	i s p o z u		m ẽj	p aj	t a t a
Portuguese	m ɐ ɾ i ð u	ɨ ʃ p o z u		m ẽj	p aj	t a t a
Catalan	m a r i t	ə s p ɔ s		m a ɾ ə	p a ɾ ə	
Walloon				m e r	p ɛ r	
Friulian	m a r iː t	s p oː 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 aː l i a	b eː s t i a	p i s k e m		a w e m	a k e l l u s / aṷ
French	o m a j			p w a s �õ		w a z o
Spanish			p e θ		a b e	
Italian		b i ʃ ʃ a	p e ʃ ʃ e			u t ʧ ɛ l l o
PortugueseBrazil		b i ʃ ɐ	p ej ʃ i		a v i	
Portuguese		b i ʃ ɐ	p ɐj ʃ ɨ		a v ɨ	
Catalan			p e ʃ		a w	u s e ʎ
Walloon		b i ɛ s		p ɛ h õ		u h ɛ
Friulian	n e m aː l	b i s	p ɛ s			u t ʧ j ɛ l
Romanian	n ə m a j e		p e ʃ t e			

Each segment gets a different number

1 f	2 eː	3 m	4 i	5 n	6 a	7 e	8 ɓ	9 ɾ	10 ɐ̯	11 j	12 œ	13 ɛ	14 u	15 l	16 r	17 x	18 o	19 ʎ	20 ʁ
21 iː	22 d	23 ɲ	24 ɔ	25 ə	26 o̯a	27 t	28 ɑ̃	29 ĩ	30 ɛ̯	31 ʧ	32 ɨ	33 ẽj̃	34 s	35 p	36 z	37 ʃ	38 ð	39 oː
40 k	41 ʦ	42 aː	43 ẽj̃	44 aj	45 θ	46 ej	47 ɐj	48 w	49 ɔ̃	50 h	51 õ	52 v	53 au̯	54 ɛ̃	55 ẽw̃	56 c			
57 ɨj	58 ẽ	59 ł	60 ɑ	61 uː	62 ɥ	63 y	64 g	65 ɨw	66 ɔj	67 ʣ	68 ɟ	69 ʤ	70 oj	71 gʷ	72 ŋ	73 ø	74 β		
75 ẹa	76 ʀ	77 ɛj	78 kʲ	79 ɪ	80 ʒ	81 ɟ	82 ɣ	83 kʷ	84 ũ	85 ɻ	86 j̃	87 œ̃	88 øː	89 ʊ					
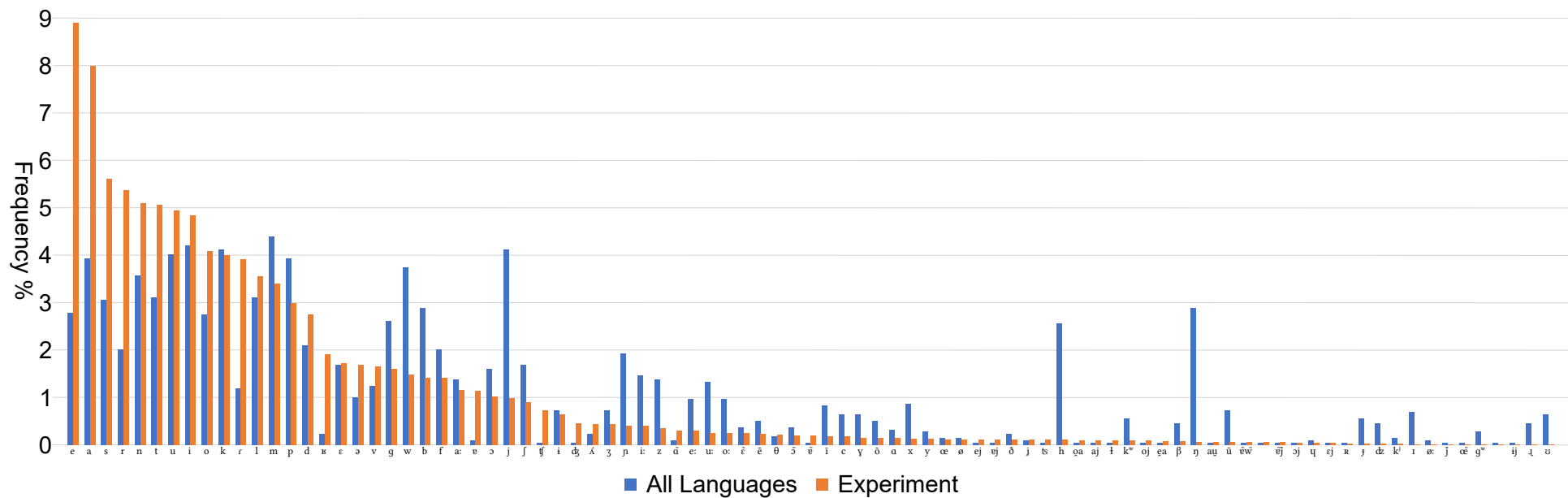
Model: "Linguistically Informed"

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

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.

1. e
2. i
3. a
4. u
5. o
6. ε
7. ø
8. Nasal Vowel
9. Vowel
10. Nasal Consonant
11. Non Sylabic Sonorant
12. Consonant
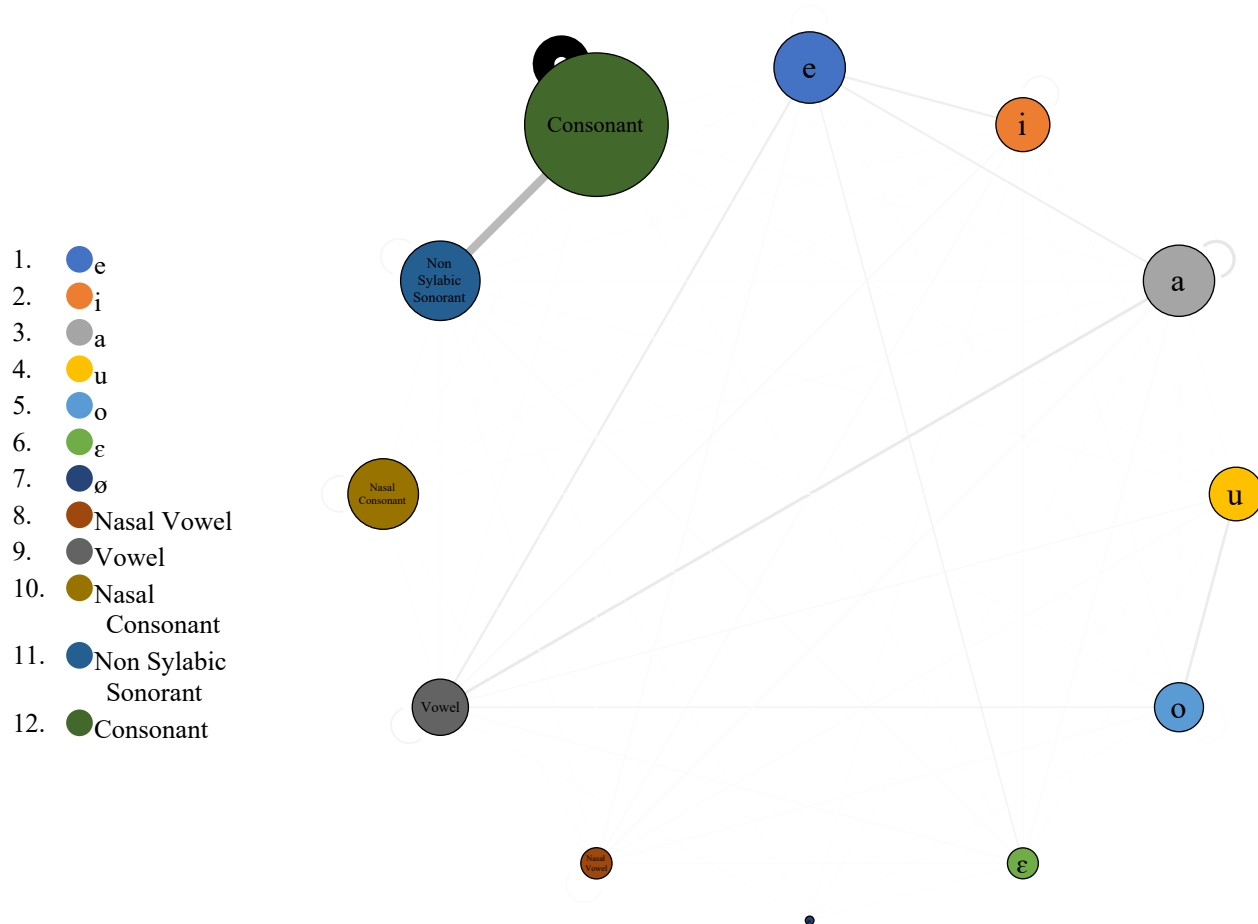
This chart shows the number of times that each segment (on the left vertical bar) differs from any other other segment in the same cognate set at the same aligned segment position. The rates of transitions to gaps and from gaps are ignored. The segments are grouped into partitions, represented by the 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.

?