

# **Languages of the World**

Dr. Asya Pereltsvaig

# Today's "secret location": Rome



# Why are these words similar across these languages yet not exactly the same?

UK Data Explorer

## European word translator

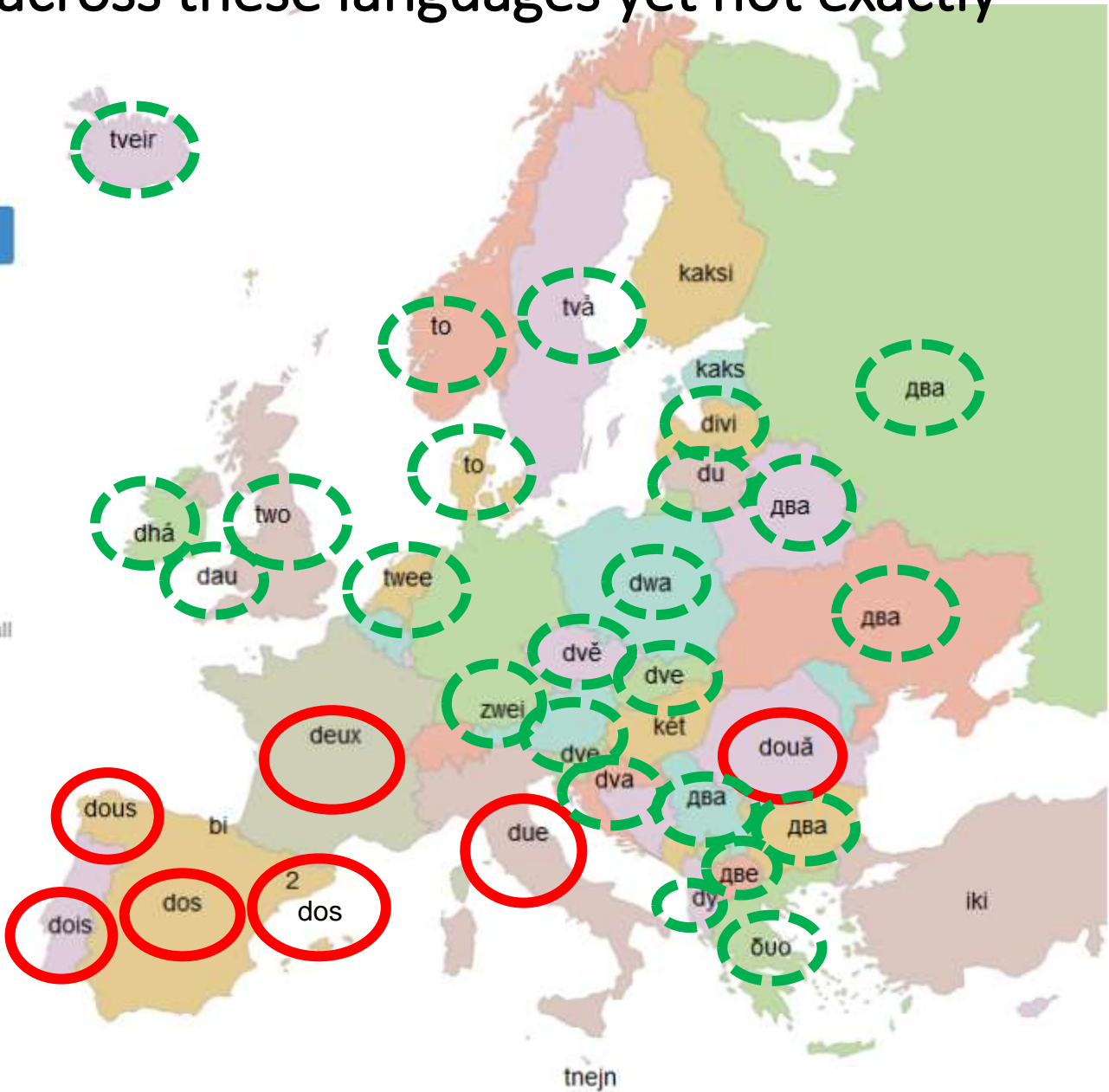
Enter one or two lower-case English words to see translations from Google Translate.

Examples: banana the cat she runs

Random words: fleet reaction

A few things to keep in mind:

- Translations are generated by Google Translate. Some may be inaccurate or use non-European (e.g. Brazilian Portuguese) words.
- Just one translation is provided for each word; watch out for words with multiple meanings!
- The Google Translate API does not yet translate into all [European languages](#).
- If Google Translate cannot find a translation, it simply shows the English word.







# Why are these words similar across these languages yet not exactly the same?

Data Explorer

## European word translator

Enter one or two lower-case English words to see translations from Google Translate.

Translate it!

Examples: banana the cat she runs

Random words: freedom supposedly

A few things to keep in mind:

- Translations are generated by Google Translate. Some may be inaccurate or use non-European (e.g. Brazilian Portuguese) words.
- Just one translation is provided for each word; watch out for words with multiple meanings!
- The Google Translate API does not yet translate into all [European languages](#).
- If Google Translate cannot find a translation, it simply shows the English word.

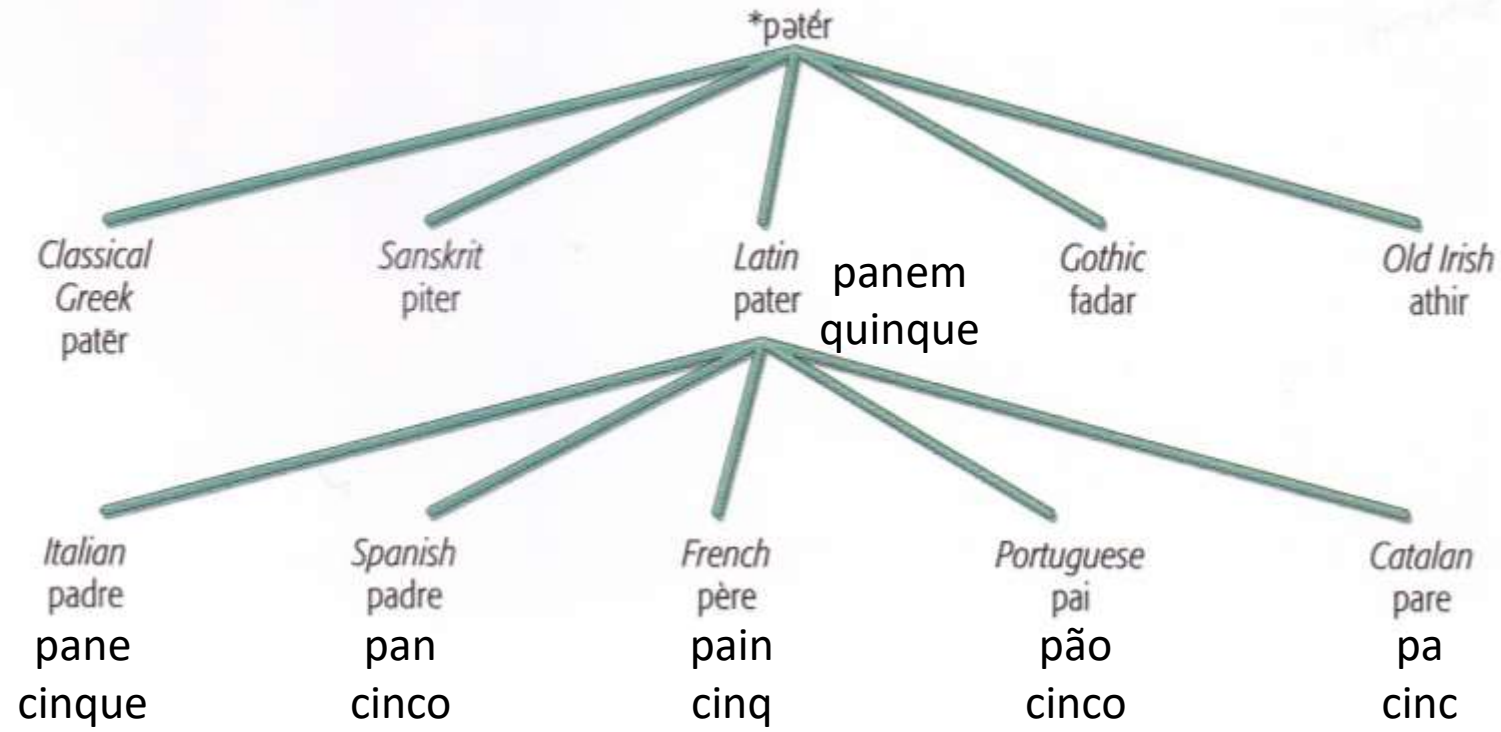




# Major Romance languages in Europe



# Reconstructing ancestral languages

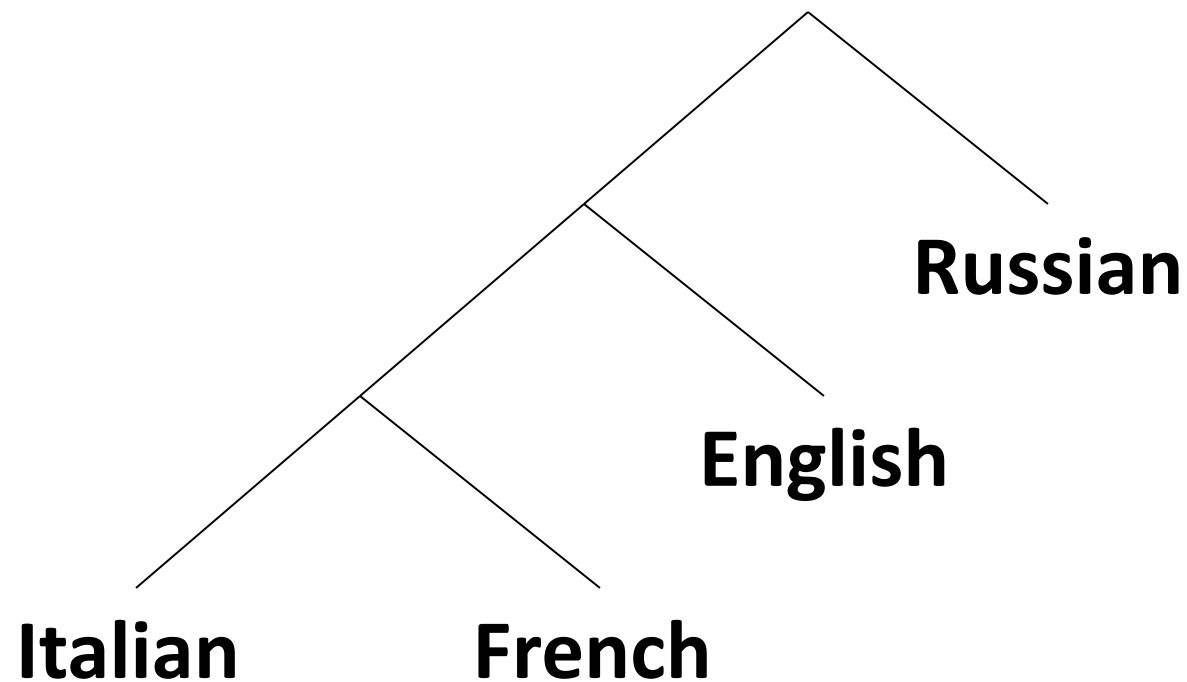




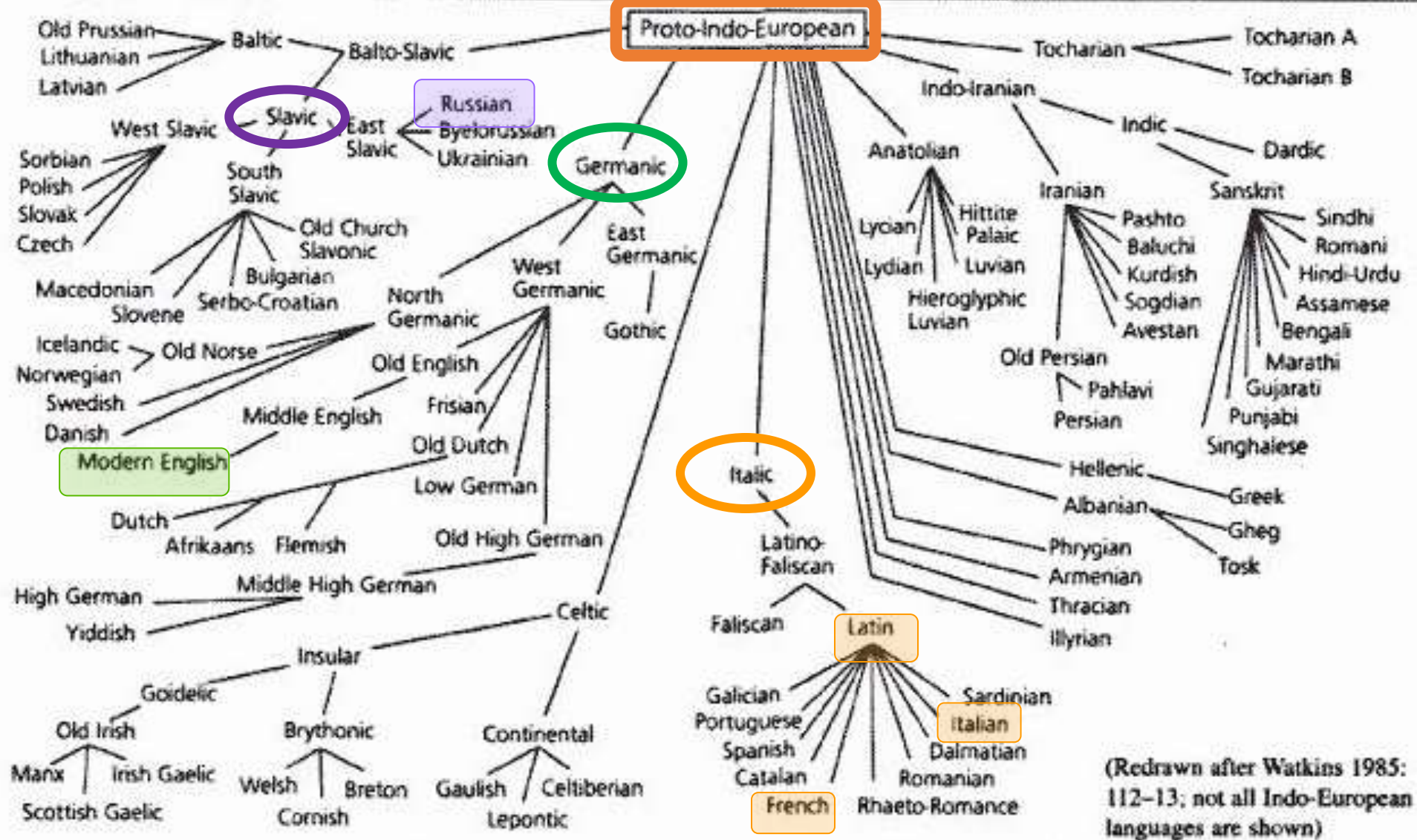
# Degrees of relatedness

<b>Italian</b>	<b>French</b>	<b>English</b>	<b>Russian</b>
sangue	sang	blood	krov'
osso	os	bone	kost'
verme	ver	worm	červ'
padre	père	father	otec
sole	soleil	sun	solnce
tre	trois	three	tri

# Building a family tree (phylogeny)



# Indo-European Language Family



(Redrawn after Watkins 1985: 112-13; not all Indo-European languages are shown)

## What language is that?

A sheep that had no wool saw horses, one of them pulling a heavy wagon, one carrying a big load, and one carrying a man quickly. The sheep said to the horses: “My heart pains me, seeing a man driving horses.” The horses said: “Listen, sheep, our hearts pain us when we see this: a man, the master, makes the wool of the sheep into a warm garment for himself. And the sheep has no wool.” Having heard this, the sheep fled into the plain.

h<sub>2</sub>áuej h<sub>1</sub>iosméj h<sub>2</sub>u|h<sub>1</sub>náh<sub>2</sub> né h<sub>1</sub>ést, só h<sub>1</sub>ékúoms derkt. só g<sup>w</sup>rh<sub>x</sub>úm uógh<sup>h</sup>om uegh<sup>h</sup>ed; só mégh<sub>2</sub>η b<sup>h</sup>órom;  
só d<sup>h</sup>gh<sup>h</sup>émonη h<sub>2</sub>óku b<sup>h</sup>ered. h<sub>2</sub>óuis h<sub>1</sub>ék<sup>w</sup>oib<sup>h</sup>ios ueuked: “d<sup>h</sup>gh<sup>h</sup>émonη spēkiōh<sub>2</sub> h<sub>1</sub>ékúoms-k<sup>w</sup>e h<sub>2</sub>ágeti,  
kér moi ag<sup>h</sup>nutor”. h<sub>1</sub>ékúōs tu ueukond: “klud<sup>h</sup>i, h<sub>2</sub>ouej! tód spēkiōmes, ηsméj ag<sup>h</sup>nutór kér: d<sup>h</sup>gh<sup>h</sup>émō, pótis,  
sē h<sub>2</sub>áuiēs h<sub>2</sub>u|h<sub>1</sub>náh<sub>2</sub> g<sup>wh</sup>érmom uéstrom uept, h<sub>2</sub>áuib<sup>h</sup>ios tu h<sub>2</sub>u|h<sub>1</sub>náh<sub>2</sub> né h<sub>1</sub>esti. tód kékluuós h<sub>2</sub>óuis  
h<sub>2</sub>agróm b<sup>h</sup>uged.

<https://soundcloud.com/archaeologymag/sheep-and-horses>

<https://www.youtube.com/watch?v=YJ-Ru9djdmU>



# Comparative Reconstruction: Lexicon



# Reconstructing ancestral languages, step 1:

**choose the right words to compare (cognates)**

Words can be similar because of:

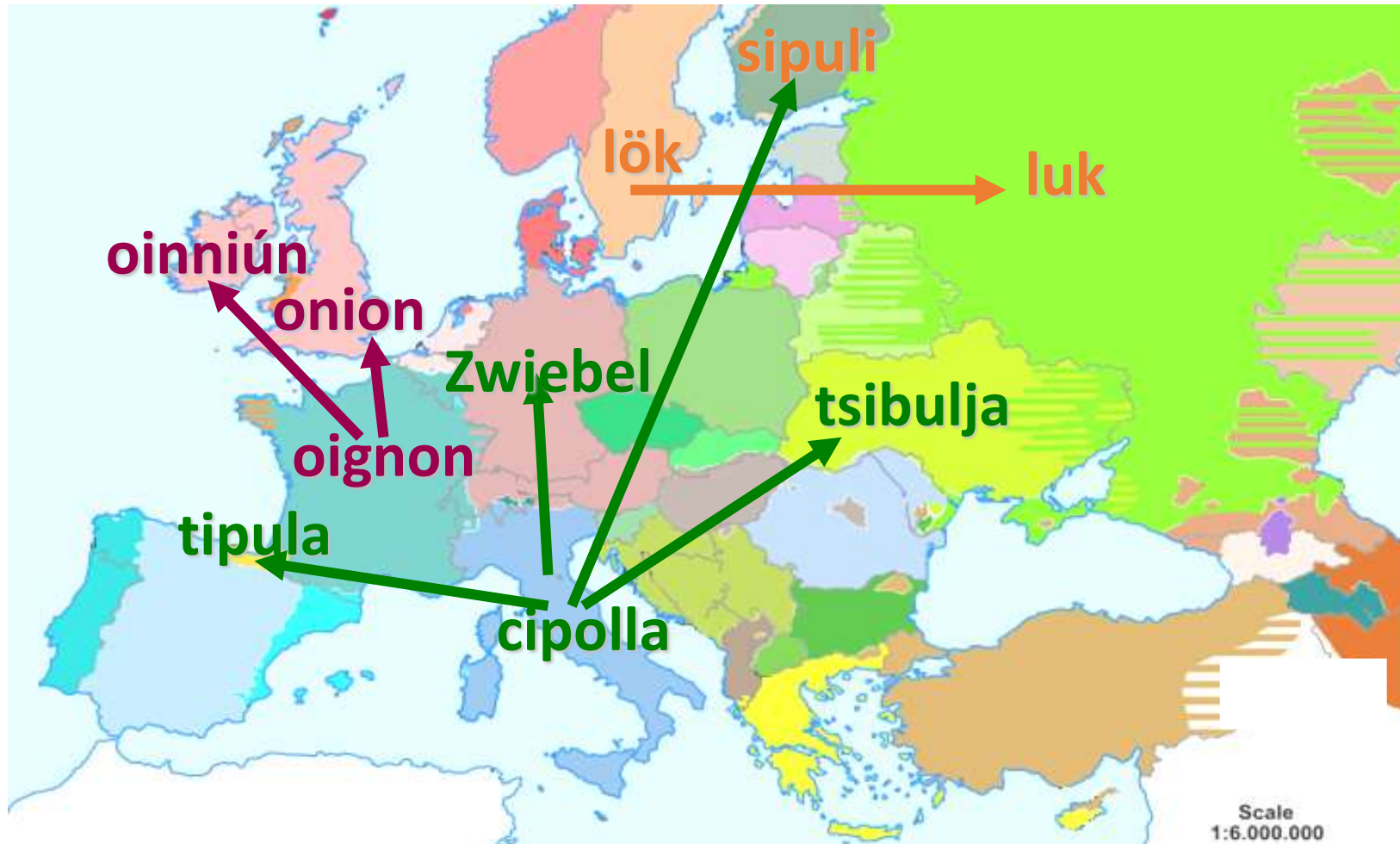
- an accident (“false cognates”)
- borrowing (“horizontal transmission”)
- common descent from the same ancestral form → **cognates**

# Words can be similar by chance

Russian <i>strannyj</i>	Italian <i>strano</i>	'strange, weird'
English <i>bad</i>	Persian <i>bad</i>	'bad'
English <i>man</i>	Korean <i>man</i>	'man'
German <i>nass</i>	Zuni <i>nas</i>	'wet'
Italian <i>donna</i>	Japanese <i>onna</i>	'lady'
Hungarian <i>fiú</i>	Romanian <i>fiú</i>	'boy, son'



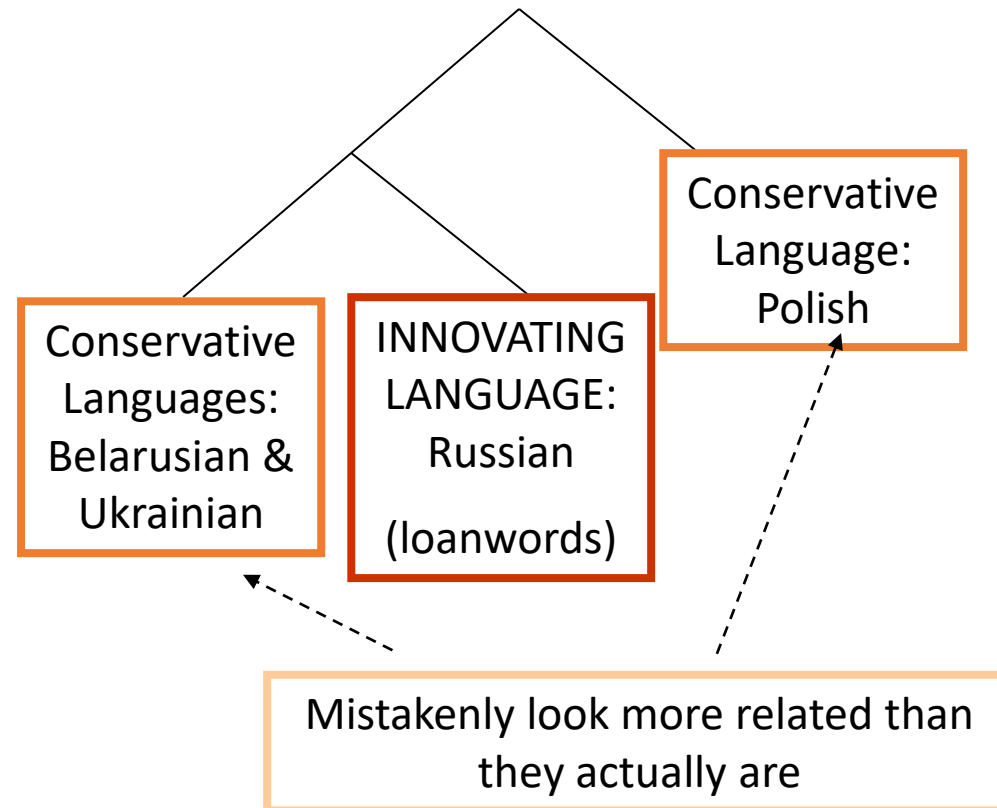
Words are easily borrowed across language family boundaries



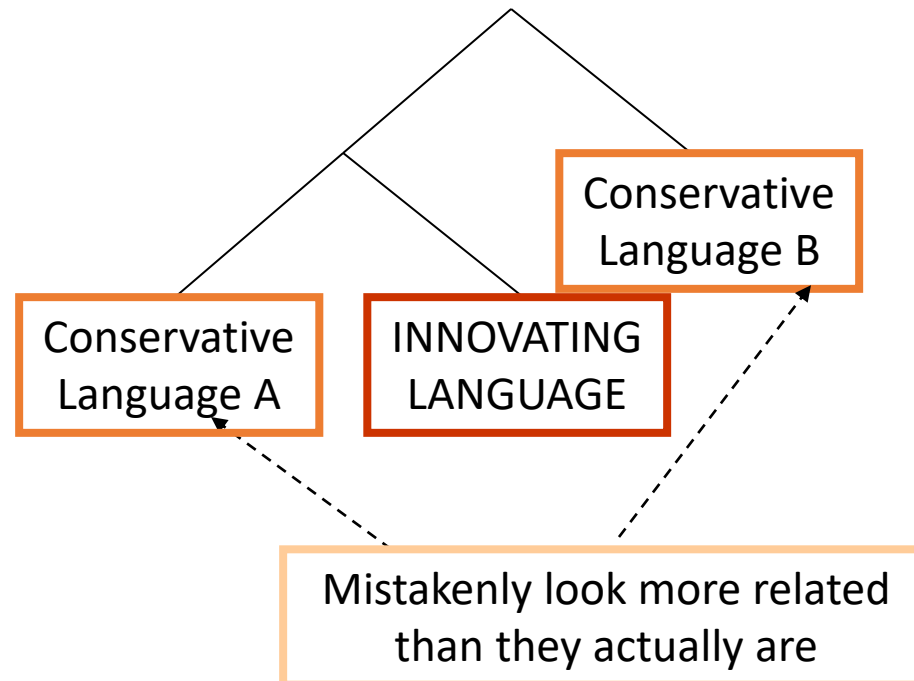
# Borrowing may distort language classification

	Russian	Belarusian	Ukrainian	Polish
January	<i>janvar'</i>	<i>studzen'</i>	<i>sičen'</i>	<i>styczeń</i>
February	<i>fevral'</i>	<i>ljuty</i>	<i>ljutyj</i>	<i>luty</i>
March	<b><i>mart</i></b>	<i>sakavik</i>	<i>berezen'</i>	<b><i>marzec</i></b>
April	<i>aprel'</i>	<i>krasavik</i>	<i>kviten'</i>	<i>kwiecień</i>
May	<b><i>maj</i></b>	<b><i>maj</i></b>	<i>traven'</i>	<b><i>maj</i></b>
June	<i>ijun'</i>	<i>červen'</i>	<i>červen'</i>	<i>czerwiec</i>
July	<i>ijul'</i>	<i>lipen'</i>	<i>lypen'</i>	<i>lipiec</i>
August	<i>avgust</i>	<i>žniven'</i>	<i>serpen'</i>	<i>sierpień</i>
September	<i>sentjabr'</i>	<i>verasen'</i>	<i>veresen'</i>	<i>wrzeszień</i>
October	<i>oktjabr'</i>	<i>kastryčnik</i>	<i>žovten'</i>	<i>październik</i>
November	<i>nojabr'</i>	<i>listapad</i>	<i>lystopad</i>	<i>listopad</i>
December	<i>dekabr'</i>	<i>snežan</i>	<i>hruden'</i>	<i>grudzień</i>

Russian innovations make Belarusian and Ukrainian look more similar to Polish than to Russian



# Innovation makes sister and cousin look more similar





# What words are most likely to be borrowed?

- A. Words for family members (kinship terms)
- B. Pronouns
- C. Numbers 1 - 10
- D. Words for plants and animals, cultural innovations, technology

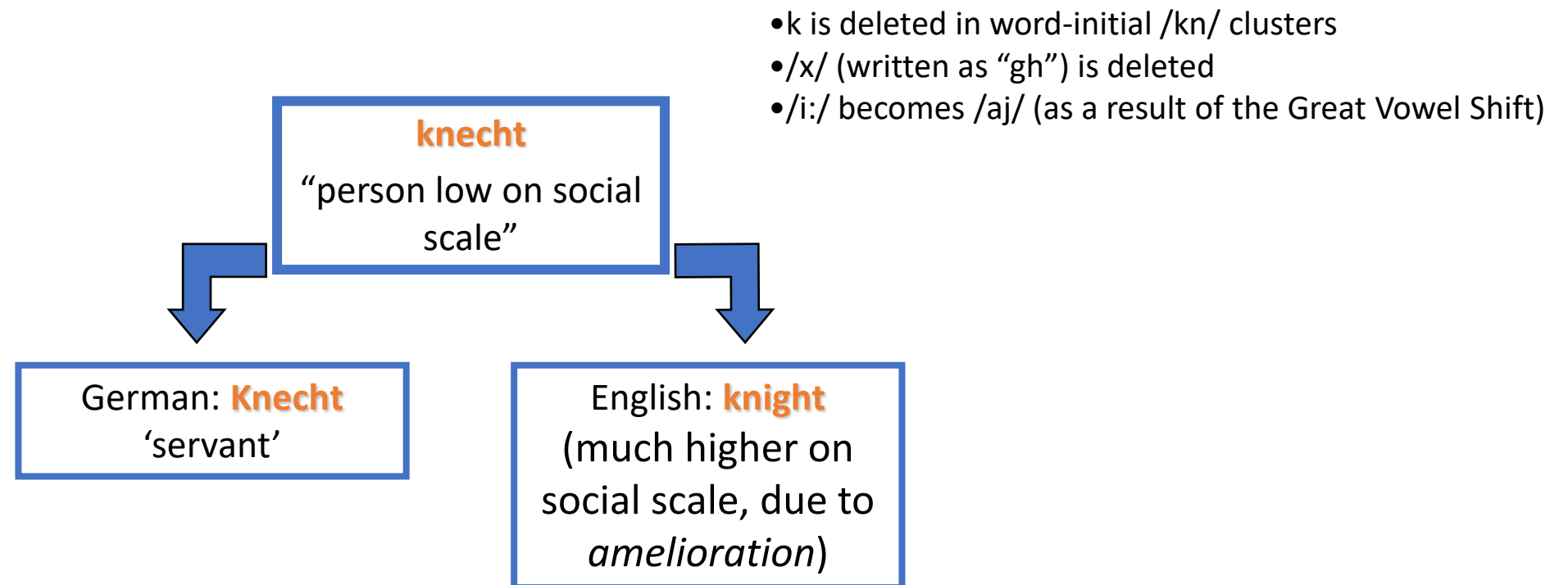


**It's D.**





Cognates = words of similar meaning and similar sound, derived from a common ancestral form



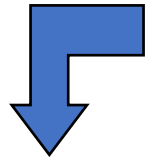


# Sometimes sound changes are not obvious

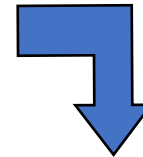
- loss of initial /f/
- /lj/ turning into /x/

Vulgar Latin *folia* /'folja/

- final /a/ becomes zero (via a schwa)
- /lj/ becomes /j/
- stressed /o/ becomes /ø/
- initial /f/ is unscathed



Spanish *HOJA* /oxa/



French *FEUILLE* /føj/

# Reconstructing ancestral languages, step 2:

**align and determine sound correspondences**

# Reconstructing ancestral forms: align and determine sound correspondences

'goat'		#1	#2	#3
Italian	/k <b>a</b> p <b>r</b> a/	k	a	p
Spanish	/k <b>a</b> b <b>r</b> a/	k	a	b
Portuguese	/k <b>a</b> b <b>r</b> a/	k	a	b
French	/ʃ <b>ε</b> v <b>r</b> ə/	ʃ	ε	v
ancestral		?	?	?

***Which sound did their common ancestor have?***

# Reconstructing ancestral languages, step 3:

**reconstructing ancestral form**

Which sound did the common ancestor have?  
And how do we decide? Majority?

'goat'		#1	#2	#3
Italian	/k <b>a</b> pra/	k	a	p
Spanish	/k <b>a</b> bra/	k	a	b
Portuguese	/k <b>a</b> bra/	k	a	b
French	/ʃ <b>ε</b> vrə/	ʃ	ε	v
ancestral (by majority)	kabr-?	k	a	<b>b</b>

*Majority often gets us the right result but not always*



# Fassano Ladin



# Majority of which languages? How do we select them?

'goat'		#1	#2	#3
Italian	/k <b>a</b> pra/	k	a	p
Spanish	/k <b>a</b> bra/	k	a	b
<b>Fassano Ladin</b>	/t <b>f</b> <b>a</b> vra/	<b>tʃ</b>	<b>a</b>	<b>v</b>
French	/ʃ <b>ε</b> vrə/	ʃ	ε	v
ancestral (by majority)	kabr-?	k	a	<b>v</b>

# Reconstructing 'goat' in Latin...

- kab**ra** or kav**ra**?
- **b** based on Italian, Spanish, Portuguese and French
- **v** based on Italian, Spanish, Fassano Ladin and French
- **Neither is correct!**

The majority principle sometimes gives the right results, but sometimes it doesn't!

'goat'		#1	#2	#3
Italian	/k <u>a</u> pra/	k	a	p
Romanian	/k <u>a</u> pra/	k	a	p
Spanish	/k <u>a</u> bra/	k	a	b
Portuguese	/k <u>a</u> bra/	k	a	b
Catalan	/k <u>a</u> bra/	k	a	b
Fassano Ladin	/t <u>f</u> avra/	tʃ	a	v
French	/ʃ <u>ε</u> vrə/	ʃ	ε	v
<b>Latin</b>	<b><i>capra</i></b>	<b><i>k</i></b>	<b><i>a</i></b>	<b><i>p</i></b>



*Some sound changes are “one way”!*

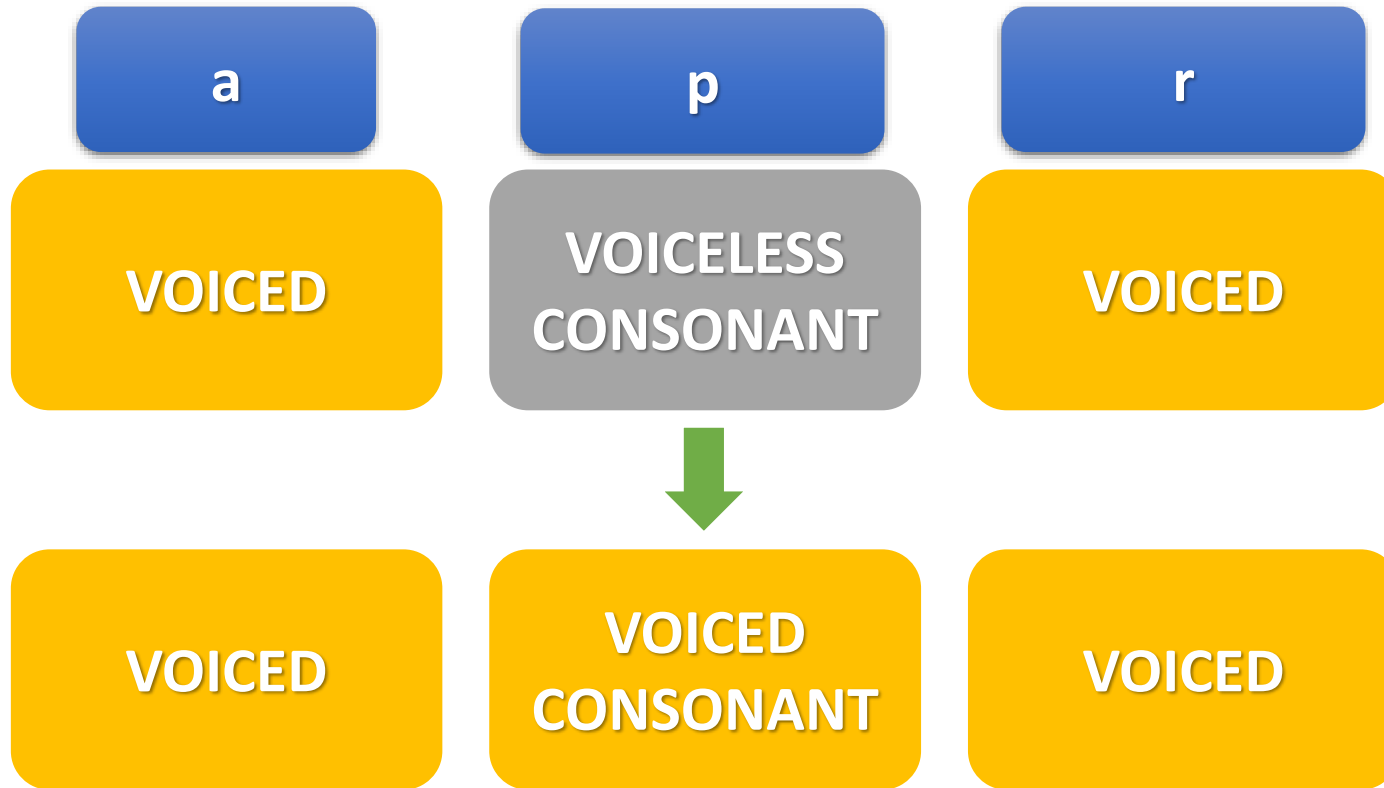
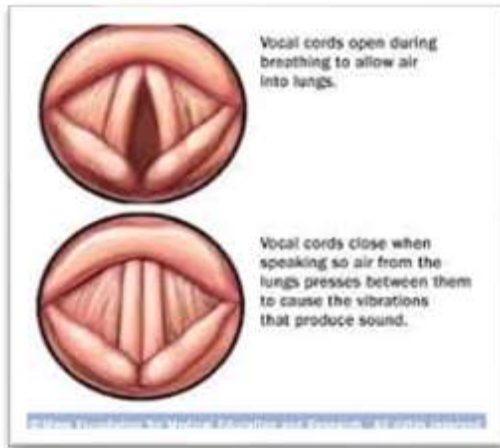
**Not CABRICORN!**

But wait, why are we reconstructing Latin? Don't we have written documents in Latin?

- Test the methodology:
  - Are we building a model of the right thing?
  - Are we building it right?
- Majority Principle isn't always right
- Some changes are attested while the reverse is not
- Some changes are common while the reverse is rare



Some sound changes are “one way”: voicing between voiced sounds



# Reconstructing the undocumented common ancestor of English, Latin Greek and Old Irish

English	Latin	Ancient Greek	Old Irish
<i>six</i>	<i>sex</i>	<i>hexa</i>	<i>se</i>
<i>seven</i>	<i>septem</i>	<i>hepta</i>	<i>seacht</i>
<i>salt</i>	<i>sal</i>	<i>hal</i>	<i>salann</i>

Which is the ancestral form: s or h?

**The ancestral form is /s/.**

**The s → h change is very common.**

**In fact, oral fricative → h is very common.**

# Some sound changes are “one way”

«heaven»	
Spanish	/θjelo/
Portuguese	/sɛw/
Catalan	/sɛl/
Occitan	/sɛl/
Italian	/tʃɛlo/
Romanian	/tʃɛr/
Sardinian	/kelu/

**What sound was in the ancestral language, based on majority?**

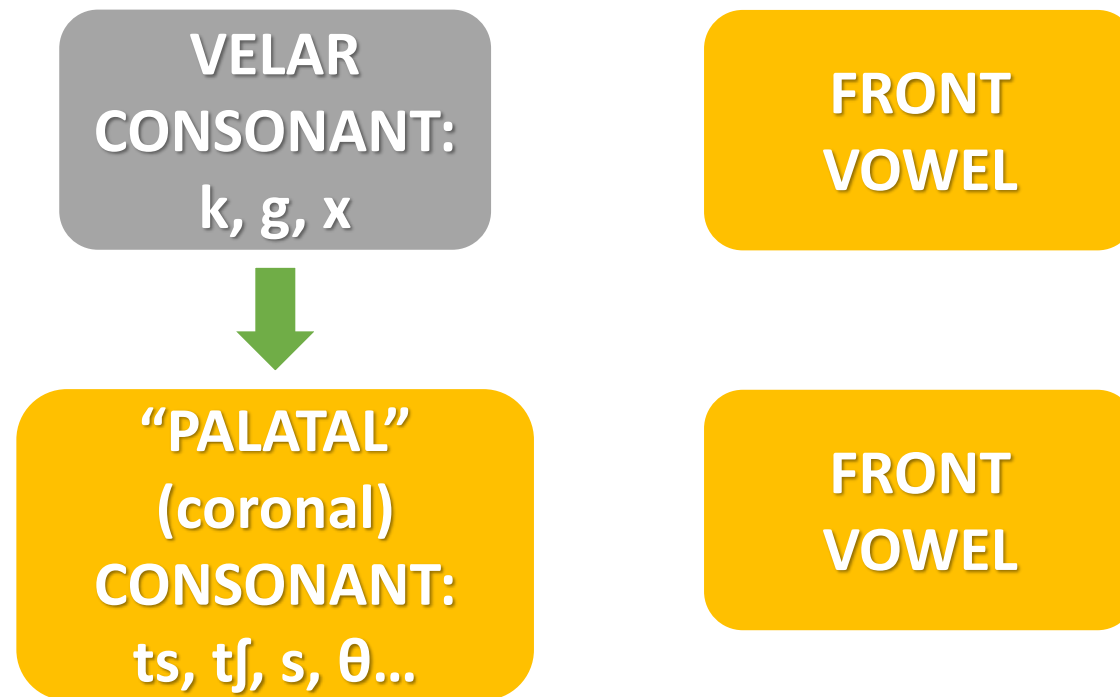
**In reality, Latin /k/ *caelum***

How do we know how Latin “c” was pronounced?



Κικερώνας  
/kikeronas/

Some sound changes are “one way”: palatalization



Old Novgorod Russian: Birch bark document №130  
(2<sup>nd</sup> half of the 14<sup>th</sup> c.)



КВНГАА:К:ЛОКОТОХЪРНБЕЗДОЛОКТИЦЪАЛН\*ТЪВКЮЛОЛАКШН:ѢА:ЛОКТИ  
ХЪРНКВАНВАСАЦВАМКШННА:ѢВ:ЛОКТИВДАМОЛЧНПОЛОТРЕТНАЛАНАТЕЛО  
КТИХЪРНЦМЪЛНТАВКЦАОЛБ:А:ЛОКТИХЪРН



# Old Novgorod Russian: Birch bark document №130 (2<sup>nd</sup> half of the 14<sup>th</sup> c.)

У ВИГАРА :К: ЛОКОТО **ХЪРИ** БЕЗО ЛОКТИ У ВАЛИТА В КЮЛОЛАКШИ :І:Д: ЛОКТИ  
**ХЪРИ** У ВАИВАСА У ВААКШИНА :ІВ: ЛОКТИ ВОДМОЛУ И ПОЛО ТРЕТИА НАЦАТЕ ЛО  
КТИ **ХЪРИ** У МЪЛИТА В КУРОЛЪ :Д: ЛОКТИ **ХЪРИ**

to Vigar 20 elbows of **XER'** without elbow to Valit in Kylolakša\* 14 elbows  
of **XER'** to Vaivas to Vayakshin 12 elbows of **VODMOL** and half 13<sup>th</sup> elbows  
of **XER'** to Melit in Kurolä\*\* 4 elbows of **XER'**

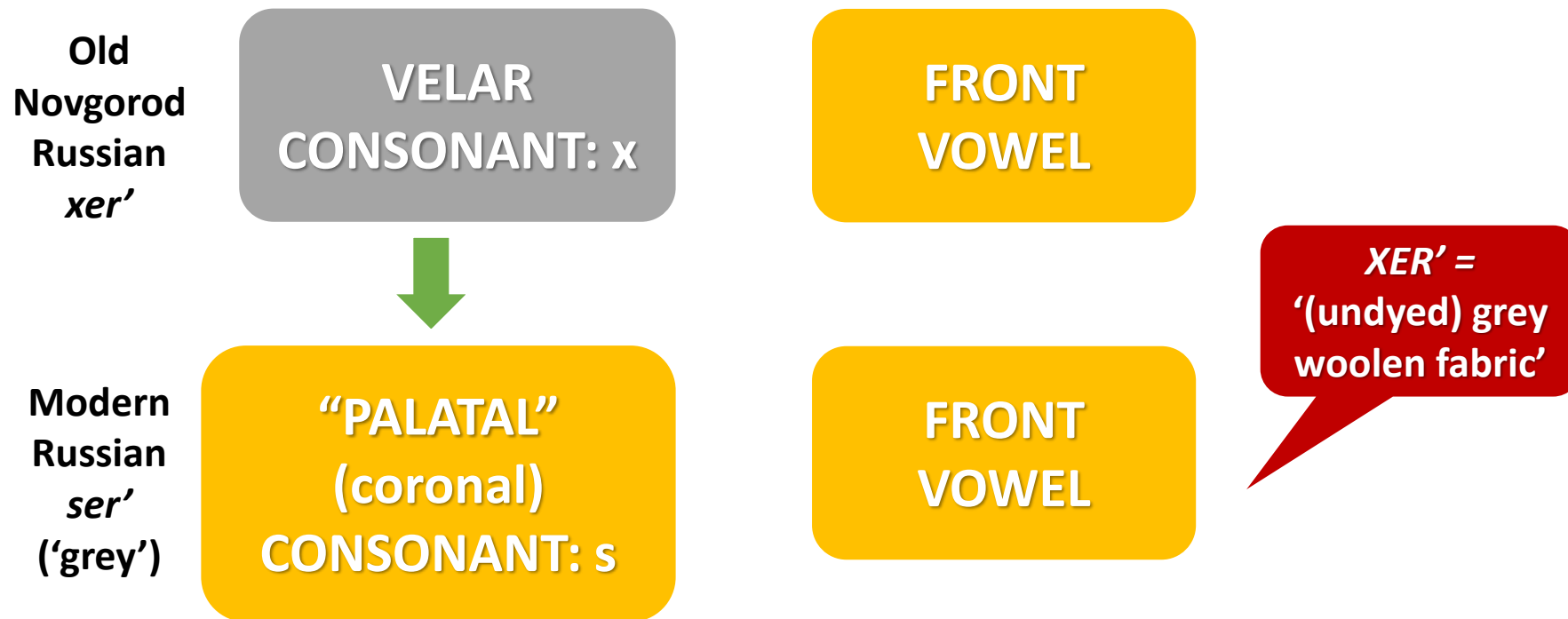
## WHAT is XER'?

\*Also mentioned in birch bark document # 248

\*\*Also mentioned in birch bark document #278



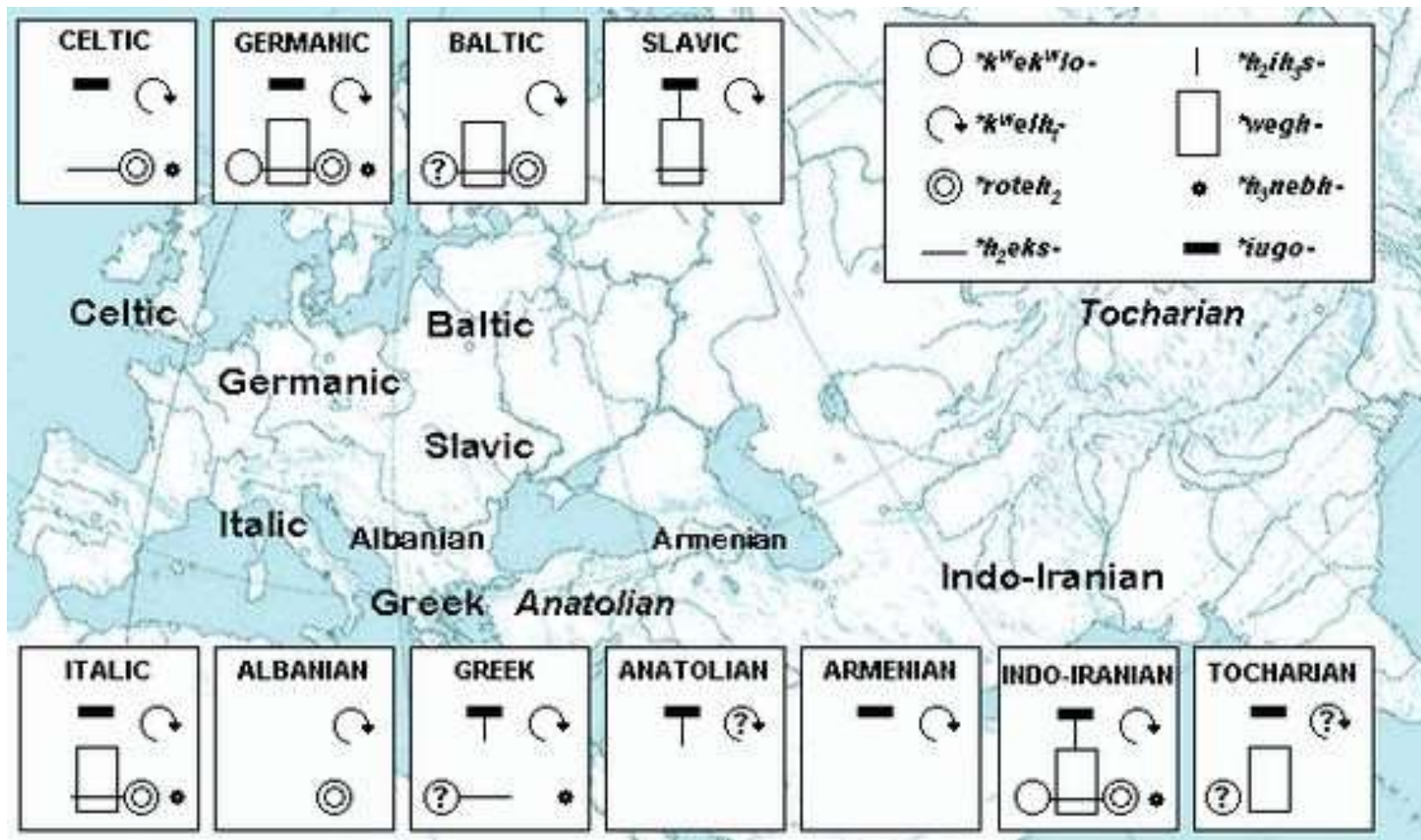
# “Second Slavic Palatalization” is lacking in Old Novgorod



So we reconstruct the ancestral language,  
then what?

- We can examine the reconstructed vocabulary for clues as to speakers' environment, lifestyle, beliefs...
- We can trace migrations, determine where a certain group comes from...
- Establish patterns of contact between groups (trade, intermarriage)

For example, words for “Horse & Wheel” can be reconstructed for PIE



# Theories of PIE Homeland: 'Steppe' or 'Kurgan' Hypothesis vs. Anatolian Hypothesis

