Languages of the World

Number of Languages

Estimates of Diversity

Source	Number of Languages
Ethnologue 1996	**6,703**
Ethnologue 2005	**6,912** (+209)
Ethnologue 2016	**7,097** (+185)

Estimates of Diversity

Source	Number of Languages
Ethnologue 1996	**6,703**
Ethnologue 2005	**6,912** (+209)
Ethnologue 2016	**7,097** (+185)

- **Not the case** that dozens of new languages are appearing every year!

Estimates of Diversity

Source	Number of Languages
Ethnologue 1996	**6,703**
Ethnologue 2005	**6,912** (+209)
Ethnologue 2016	**7,097** (+185)

- **Not the case** that dozens of new languages are appearing every year!

Two real reasons:
1. **Records become more complete**

Estimates of Diversity

Source	Number of Languages
Ethnologue 1996	**6,703**
Ethnologue 2005	**6,912** (+209)
Ethnologue 2016	**7,097** (+185)

- **Not the case** that dozens of new languages are appearing every year!

Two real reasons:
1. **Records become more complete**
2. **Different decisions are made about what constitutes a "language"**

Two Languages or Two Dialects?

1. People think of them as the same language

Two Languages or Two Dialects?

1. **People think of them as the same language**
    - **Yeah, and you're gonna to trust people's intuitions?**

Two Languages or Two Dialects?

1. **People think of them as the same language**
2. **They are mutually intelligible**

Two Languages or Two Dialects?

1. **People think of them as the same language**

2. **They are mutually intelligible**

   - **Gradient and hard to measure**

Two Languages or Two Dialects?

1. **People think of them as the same language**

2. **They are mutually intelligible**

   - **Gradient and hard to measure**

   Example: <u>Newfoundland English</u>

   What percent did you understand on the first listen?

Two Languages or Two Dialects?

1. People think of them as the same language
2. They are mutually intelligible
3. An official body says they are the same

Two Languages or Two Dialects?

1. People think of them as the same language
2. They are mutually intelligible
3. An official body says they are the same
   - Blatant political biases?
   - "A language is a dialect with an army and a navy" - M. Weinreich

Serbo-Croatian: Definition 1 ✗   Definition 2 ✔

- **Western South-Slavic languages/dialects**

- **Informally divided into Bosnian, Croatian, and Serbian, etc.**

- **Actual divisions do not follow national boundaries**

DIALECTS OF THE WESTERN BRANCH
OF SOUTH-SLAVIC LANGUAGES

STOKAVIAN
- ekavian
- ijekavian
- ikavian, partially ikavian

KAJKAVIAN
- ekavian
- ijekavian/ekavian

CAKAVIAN
- ekavian
- ekavian/ikavian
- ikavian
- ijekavian

Serbo-Croatian: Definition 1 ✗   Definition 2 ✓

Bosnian

Sva ljudska bića rađaju se slobodna i jednaka u dostojanstvu i pravima. Ona su obdarena razumom i sviješću i treba da jedno prema drugome postupaju u duhu bratstva. (*UDHR Article I*)

Croatian
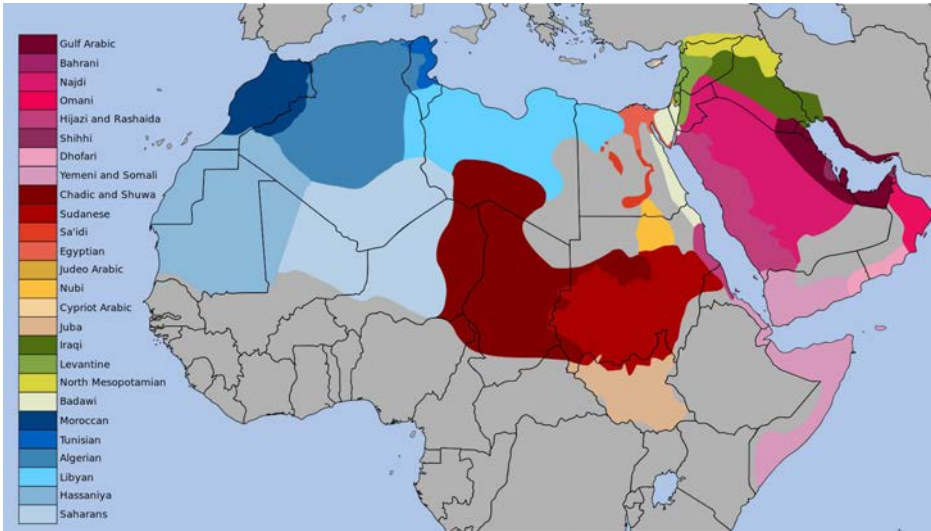
Sva ljudska bića rađaju se slobodna i jednaka u dostojanstvu i pravima. Ona su obdarena razumom i sviješću i trebaju jedna prema drugima postupati u duhu bratstva.

Serbian

Sva ljudska bića rađaju se slobodna i jednaka u dostojanstvu i pravima. Ona su obdarena razumom i svešću i treba jedni prema drugima da postupaju u duhu bratstva.

Serbo-Croatian: Definition 1 ✗   Definition 2 ✓

Bosnian

Sva ljudska bića rađaju se slobodna i jednaka u dostojanstvu i pravima. Ona su obdarena razumom i sv**ij**ešću i treba **da** jedn**o** prema drug**ome** postupa**ju** u duhu bratstva. (*UDHR Article I*)

Croatian

Sva ljudska bića rađaju se slobodna i jednaka u dostojanstvu i pravima. Ona su obdarena razumom i sv**ij**ešću i treba**ju** jedn**a** prema drug**ima** postupa**ti** u duhu bratstva.

Serbian

Sva ljudska bića rađaju se slobodna i jednaka u dostojanstvu i pravima. Ona su obdarena razumom i svešću i treba jedn**i** prema drug**ima** **da** postupa**ju** u duhu bratstva.

Arabic: Definition 1 ✓   Definition 2 ✗

Gulf Arabic
Bahrani
Najdi
Omani
Hijazi and Rashaida
Shihhi
Dhofari
Yemeni and Somali
Chadic and Shuwa
Sudanese
Sa'idi
Egyptian
Judeo Arabic
Nubi
Cypriot Arabic
Juba
Iraqi
Levantine
North Mesopotamian
Badawi
Moroccan
Tunisian
Algerian
Libyan
Hassaniya
Saharans

- **Spoken across North Africa and the Middle East**
- **Descendents of 8th Century Classical Arabic**
- **Consider Modern Standard Arabic, Cairene (NE Africa), and Maghrebi (NW Africa)**
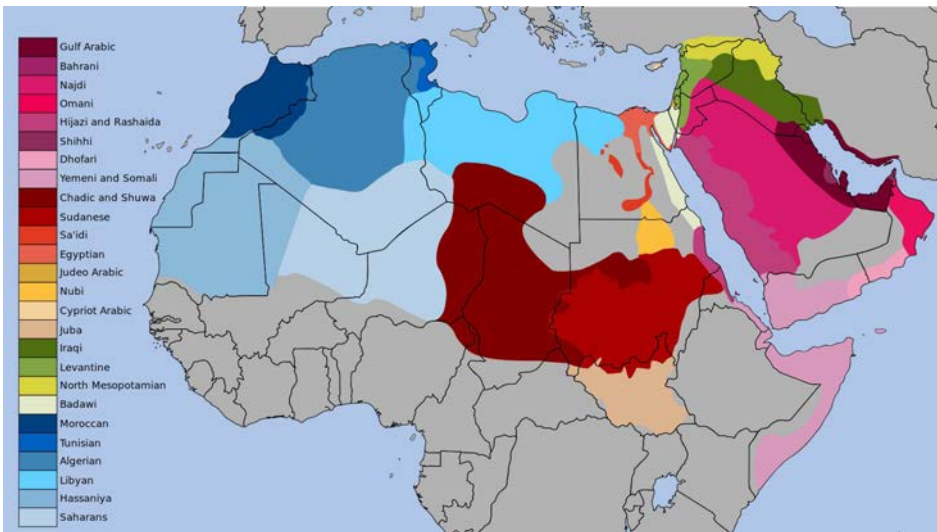
Arabic: Definition 1 ✓   Definition 2 ✗

- **Spoken across North Africa and the Middle East**
- **Descendents of 8th Century Classical Arabic**
- **Consider Modern Standard Arabic, Cairene (NE Africa), and Maghrebi (NW Africa)**

All Arabic varieties are traditionally considered a single language, but speakers are aware of mutual intelligibility.

Arabic: Definition 1 ✓   Definition 2 ✗

Modern Standard

Yūladu jamī'u n-nāsi aḥrāran mutasāwīna fī l-karāmati wa-l-ḥuqūq. Wa-qad wuhibū 'aqlan wa-ḍamīran wa-'alayhim an yu'āmila ba'ḍuhum ba'ḍan bi-rūḥi l-ikhā'. (*UDHR Article I*)

Cairene

'il-madda 'il-'awwalaniyya 'il-bani'admiin kulluhum mawluudiin ħurriin wi mitsawwyiin fil-karama wil-ħu'uu'. 'itwahab-luhum 'il-'a'l wiḍ-ḍamiir wil-mafruuḍ yi'amlu ba'ḍ bi-ruuħ 'il-'uxuwiyya.
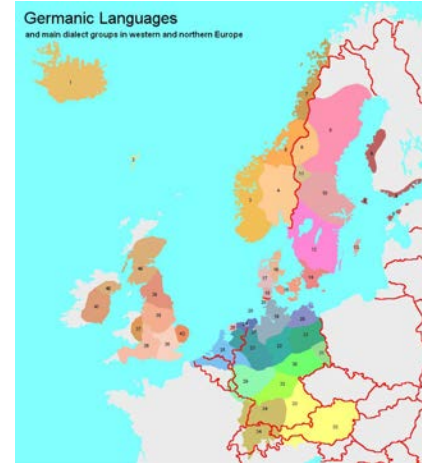
Maghrebi

kulluu nafar fii wilaada huur semsem fii karaama we semsem fii hokuuk. Alla fii 'i'tii huwwa 'akel we damiir, we laazim huwwa sawwi ma' taani nafar semsem aax.

Arabic: Definition 1 ✓ Definition 2 ✗

Modern Standard

Yūladu jamī'u **n**-nāsi aḥrāran **mutasāwīna fī** l-**karāmati** wa-l-ḥuqūq. Wa-qad wuhibū 'aqlan wa-ḍamīran wa-'alayhim an yu'āmila ba'ḍuhum ba'ḍan bi-rūḥi l-ikhā'. (*UDHR Article I*)

Cairene

'il-madda **'il**-'awwalaniyya **'il**-bani'admiin **kulluhum mawluudiin ħurriin wi mitsawwyiin fi**l-**karama wil**-ħu'uu'. 'itwahab-luhum 'il-'a'l wiḍ-ḍamiir wil-mafruuḍ yi'amlu ba'ḍ bi-ruuħ 'il-'uxuwiyya.

Maghrebi

kulluu nafar **fii wilaada huur** semsem **fii karaama we** semsem **fii hokuuk**. Alla fii 'i'tii huwwa 'akel we damiir, we laazim huwwa sawwi ma' taani nafar semsem aax.

Germanic:

- **West Germanic spoken in Switzerland, Austria, Germany, the Netherlands, Belgium, UK, etc.**

- **A continuum of varieties**

- **Consider Standard Dutch, Standard German, and Low German**

Germanic Languages
and main dialect groups in western and northern Europe

West Germanic: Definition 3 ?

Standard Dutch

Alle mensen worden vrij en gelijk in waardigheid en rechten geboren. Zij zijn begiftigd met verstand en geweten, en behoren zich jegens elkander in een geest van broederschap te gedragen. (*UDHR Article I*)

Standard German

Alle Menschen sind frei und gleich an Würde und Rechten geboren. Sie sind mit Vernunft und Gewissen begabt und sollen einander im Geist der Brüderlichkeit begegnen.

Low German

All de Minschen sünd frie un gliek an Wüürd un Rechten baren. Se hebbt Vernunft un een Geweten un se schüllt sik Bröder sien.

West Germanic: Definition 3 ?

Standard Dutch

Alle mensen worden vrij en gelijk in waardigheid en rechten geboren. Zij zijn begiftigd met verstand en geweten, en behoren zich jegens elkander in een geest van broederschap te gedragen. (*UDHR Article I*)

Standard German

Alle Menschen sind frei und gleich an Würde und Rechten geboren. Sie sind mit Vernunft und Gewissen begabt und sollen einander im Geist der Brüderlichkeit begegnen.

Low German

All de **Minschen sünd** frie un **gliek an Wüürd** un Rechten baren. Se **hebbt Vernunft** un een **Geweten** un se **schüllt sik** Bröder sien.

Counting Native Speakers

Most Languages have Few Speakers

- **Many languages have hundreds
  or thousand of speakers**
- **And are spoken in one area**
- **Some are at risk of dying out**

Most Languages have Few Speakers

- **Many languages have hundreds or thousand of speakers**
- **And are spoken in one area**
- **Some are at risk of dying out**

- **Only a few have millions of speakers**
- **Only a few are spoken over wide areas**
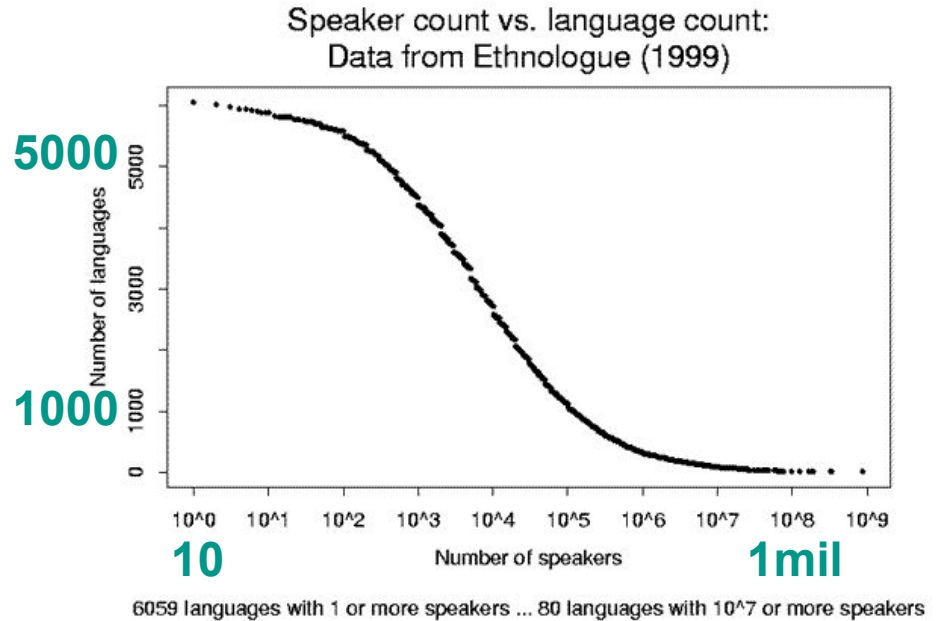
Most Languages have Few Speakers

- **Many languages have hundreds or thousand of speakers**
- **And are spoken in one area**
- **Some are at risk of dying out**

- **Only a few have millions of speakers**
- **Only a few are spoken over wide areas**

Speaker count vs. language count:
Data from Ethnologue (1999)

5000

1000

10

1mil

Number of languages

Number of speakers

10^0 10^1 10^2 10^3 10^4 10^5 10^6 10^7 10^8 10^9

6059 languages with 1 or more speakers ... 80 languages with 10^7 or more speakers

Most Spoken Languages (by native speakers)

2016 Rank	Language	Speakers (mil)
1	Mandarin	889
2	Spanish	427
3	English	339
4	Arabic (all)	267
5	Hindi	260
6	Portuguese	202
7	Bengali	189
8	Russian	171
9	Japanese	128
10	Lahnda (all)	117

2016 Rank	Language	Speakers (mil)
11	Javanese	84
12	Wu	80
13	Korean	77
14	German	77
15	French	76
16	Telugu	74
17	Marathi	72
18	Turkish	71
19	Urdu	69
20	Vietnamese	68

Most Spoken Languages (by native speakers)

2016 Rank	Language	Speakers (mil)
1	Mandarin	889
2	Spanish	427
3	English	339
4	Arabic (all)	267
5	Hindi	260
6	Portuguese	202
7	Bengali	189
8	Russian	171
9	Japanese	128
10	Lahnda (all)	117

2016 Rank	Language	Speakers (mil)
11	Javanese	84
12	Wu	80
13	Korean	77
14	German	77
15	French	76
16	Telugu	74
17	Marathi	72
18	Turkish	71
19	Urdu	69
20	Vietnamese	68

Most Spoken Languages (by native speakers)

2016 Rank	Language	Speakers (mil)
1	Mandarin	889
2	Spanish	427
3	English	339
4	Arabic (all)	267
5	Hindi	260
6	Portuguese	202
7	Bengali	189
8	Russian	171
9	Japanese	128
10	Lahnda (all)	117

2016 Rank	Language	Speakers (mil)
11	Javanese	84
12	Wu	80
13	Korean	77
14	German	77
15	French	76
16	Telugu	74
17	Marathi	72
18	Turkish	71
19	Urdu	69
20	Vietnamese	68

Most Spoken Languages (by native speakers)

2016 Rank	Language	Speakers (mil)
1	Mandarin	889
2	Spanish	427
3	English	339
4	Arabic (all)	267
5	Hindi	260
6	Portuguese	202
7	Bengali	189
8	Russian	171
9	Japanese	128
10	Lahnda (all)	117

2016 Rank	Language	Speakers (mil)
11	Javanese	84
12	Wu	80
13	Korean	77
14	German	77
15	French	76
16	Telugu	74
17	Marathi	72
18	Turkish	71
19	Urdu	69
20	Vietnamese	68

Most Spoken Languages (by native speakers)

2016 Rank	Language	Speakers (mil)
1	Mandarin	889
2	Spanish	427
3	English	339
4	Arabic (all)	267
5	Hindi	260
6	Portuguese	202
7	Bengali	189
8	Russian	171
9	Japanese	128
10	Lahnda (all)	117

2016 Rank	Language	Speakers (mil)
11	Javanese	84
12	Wu	80
13	Korean	77
14	German	77
15	French	76
16	Telugu	74
17	Marathi	72
18	Turkish	71
19	Urdu	69
20	Vietnamese	68

Who do we Count?

- **The previous counts are for native (L1) speakers only**

Who do we Count?

- **The previous counts are for native (L1) speakers only**
- **Could count second language (L2) speakers too**

Who do we Count?

- **The previous counts are for native (L1) speakers only**
- **Could count second language (L2) speakers too**
  - **How proficient does one need to be?**
  - **How to survey?**

Adding in L2 makes a huge difference

	English	French	Indonesian	Swahili
L1	339	76	23	15
L1 + L2				

Adding in L2 makes a huge difference

	English	French	Indonesian	Swahili
L1	339	76	23	15
L1 + L2	943	229	198	98

Adding in L2 makes a huge difference

	English	French	Indonesian	Swahili
L1	339	76	23	15
L1 + L2	943	229	198	98

Each is a cross-cultural trade language, a Lingua Franca

Language Death

Estimates of Diversity (revisited)

Source	Number of Languages
Ethnologue 1996	**6,703**
Ethnologue 2005	**6,912** (+209)
[Ethnologue 2016](#)	**7,097** (+185)

- The actual numbers are *declining* rapidly
- Up to one every other week

Estimates of Diversity (revisited)

Source	Number of Languages
Ethnologue 1996	**6,703**
Ethnologue 2005	**6,912** (+209)
[Ethnologue 2016](#)	**7,097** (+185)

- **The actual numbers are *declining* rapidly**
- **Up to one every other week**
- **A moribund language is one that's in danger**

Estimates of Diversity (revisited)

Source	Number of Languages
Ethnologue 1996	**6,703**
Ethnologue 2005	**6,912** (+209)
Ethnologue 2016	**7,097** (+185)

- **The actual numbers are *declining* rapidly**
- **Up to one every other week**
- **A moribund language is one that's in danger**
  - **Low population**
  - **Not being passed on normally to the younger generation**

Why Languages Die

1. Death by accident

Why Languages Die

1. **Death by accident**
    - **Maybe a storm or disease kills everyone**
    - **Remember that many languages have only tens of speakers**

Why Languages Die

1. **Death by accident**
2. **Death by assimilation**

Why Languages Die

1. **Death by accident**

2. **Death by assimilation**
   - Sometimes cultures grow to prefer the local Lingua Franca
   - Parents speak that rather than the local language to their children
   - The local language dies out

Why Languages Die

1. **Death by accident**
2. **Death by assimilation**
   - **Latin killed Etruscan, Oscan, Umbrian, etc.**
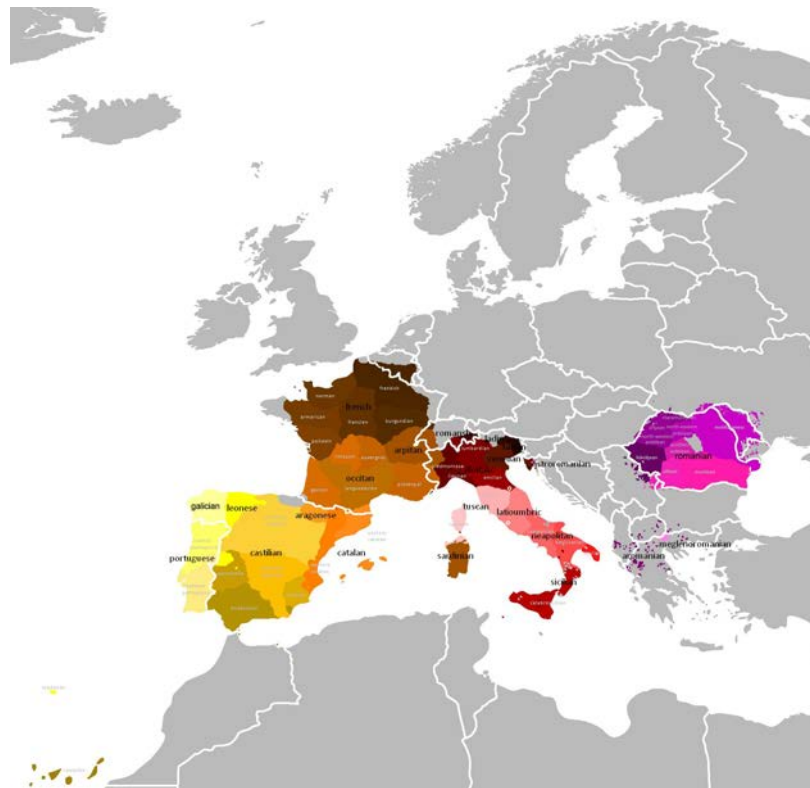

Linguistic Diversity in Pre-Roman Italy

Why Languages Die

1. **Death by accident**
2. **Death by assimilation**
3. **Death on purpose**

Why Languages Die

1. **Death by accident**
2. **Death by assimilation**
3. **Death on purpose**
   - **Some governments institute policies to suppress local languages**

Why Languages Die

1. **Death by accident**
2. **Death by assimilation**
3. **Death on purpose**
   - **Some governments institute policies to suppress local languages**
   - **Happened in the US and Australia**

Why Languages Die

1. **Death by accident**
2. **Death by assimilation**
3. **Death on purpose**

- **Some governments institute policies to suppress local languages**
- **Happened in the US and Australia**
- **Happening in France still**

PARLEZ FRANCAIS   SOYEZ PROPRES

SPEAK FRENCH   BE CLEAN

Official policy discourages the use of Occitan. Older generations rarely speak it to their children or to outsiders.

Why Languages Die

1. **Death by accident**
2. **Death by assimilation**
3. **Death on purpose**
4. **Evolution**

Why Languages Die

1. **Death by accident**
2. **Death by assimilation**
3. **Death on purpose**
4. **Evolution**
   - **Not all languages hit a dead end**
   - **Any language that stays alive changes over the generations**

Why Languages Die

1. **Death by accident**
2. **Death by assimilation**
3. **Death on purpose**
4. **Evolution**
   - **Not all languages hit a dead end**
   - **Any language that stays alive changes over the generations**
   - **Latin yielded the Romance languages, but there was no single point when Latin died.**

Latin & Romance Language Comparison

Latin   (Ea) semper antequam cenat fenestram claudit

Catalan   (Ella) sempre tanca/clou la finestra abans de sopar

Emilian   (Le) la sèra sèmpar sù la fnèstra prima ad snàr

French   Elle ferme toujours la fenêtre avant de dîner/souper

Italian   (Ella/Lei) chiude sempre la finestra prima di cenare

Occitan   (Ela) barra/tanca sempre/totjorn la fenèstra abans de sopar

Portuguese   (Ela) fecha sempre a janela antes de jantar.

Romanian   Ea închide întotdeauna fereastra înainte de cină

Sardinian   Issa serrat sempri/semper sa ventana innantis de cenài/chenàre

Spanish   (Ella) siempre cierra la ventana antes de cenar

Latin & Romance Language Comparison

Latin   (Ea) semper antequam cenat fenestram claudit

Catalan   (Ella) sempre tanca/clou la finestra abans de sopar

Emilian   (Le) la sèra sèmpar sù la fnèstra prima ad snàr

French   Elle ferme toujours la fenêtre avant de dîner/souper

Italian   (Ella/Lei) chiude sempre la finestra prima di cenare

Occitan   (Ela) barra/tanca sempre/totjorn la fenèstra abans de sopar

Portuguese (Ela) fecha sempre a janela antes de jantar.

Romanian Ea închide întotdeauna fereastra înainte de cină

Sardinian Issa serrat sempri/semper sa ventana innantis de cenài/chenàre

Spanish   (Ella) siempre cierra la ventana antes de cenar

Grammatical Diversity

Typology

- **The classification of grammatical features across language**

Typology

- **The classification of grammatical features across language**
- **Correlated with but independent of language families**

Typology

- **The classification of grammatical features across language**
- **Correlated with but independent of language families**
- **Examples features ([see more here](#)):**
  - **phonemic inventories - does the language have /q/?**
  - **word orders - Are declarative sentences SVO or SOV?**
  - **inflection - do verbs agree morphologically with their subjects?**

Typology

- **The classification of grammatical features across language**
- **Correlated with but independent of language families**
- **Examples features ([see more here](#)):**
  - **phonemic inventories - does the language have /q/?**
  - **word orders - Are declarative sentences SVO or SOV?**
  - **inflection - do verbs agree morphologically with their subjects?**
  - **Does the language express noun classes?**

Noun Classes - Grammatical Gender

- **Some languages (Spanish, Hebrew, etc.) divide nouns into masculine and feminine**

Noun Classes - Grammatical Gender

- **Some languages (Spanish, Hebrew, etc.) divide nouns into masculine and feminine**
- **These largely correlate with "natural" gender ('boy' M, 'girl' F, etc)**

Noun Classes - Grammatical Gender

- **Some languages (Spanish, Hebrew, etc.) divide nouns into masculine and feminine**
- **These largely correlate with "natural" gender ('boy' M, 'girl' F, etc)**

	F	M
boy		niño
girl	niña	
cow	vaca	
bull		toro

Noun Classes - Grammatical Gender

- **Some languages (Spanish, Hebrew, etc.) divide nouns into masculine and feminine**
- **These largely correlate with "natural" gender ('boy' M, 'girl' F, etc)**
- **But inanimate objects have gender too**
- **grammatical gender ≠ natural gender**

	F	M
boy		niño
girl	niña	
cow	vaca	
bull		toro
brick		ladrillo
table	mesa	

Noun Classes - 3 Genders

- **Some languages have a 3 gender feminine-masculine-neuter system (eg Latin, German)**

Noun Classes - 3 Genders

- **Some languages have a 3 gender feminine-masculine-neuter system (eg Latin, German)**
- **Maybe these correspond to natural gender better than M-F systems?**

Noun Classes - 3 Genders

- **Some languages have a 3 gender feminine-masculine-neuter system (eg Latin, German)**
- **Maybe these correspond to natural gender better than M-F systems?**
- **Nope…**
- **Take German, for example:**

	F	**M**	**N**
logic	die Logik		
honey		der Honig	
maiden			das Mädchen

Noun Classes - 2 Genders (revisited)

- **M-F isn't the only possible two gender system.**

Noun Classes - 2 Genders (revisited)

- **M-F isn't the only possible two gender system.**
- **Many European languages (Dutch, Swedish, etc) make a common-neuter distinction.**
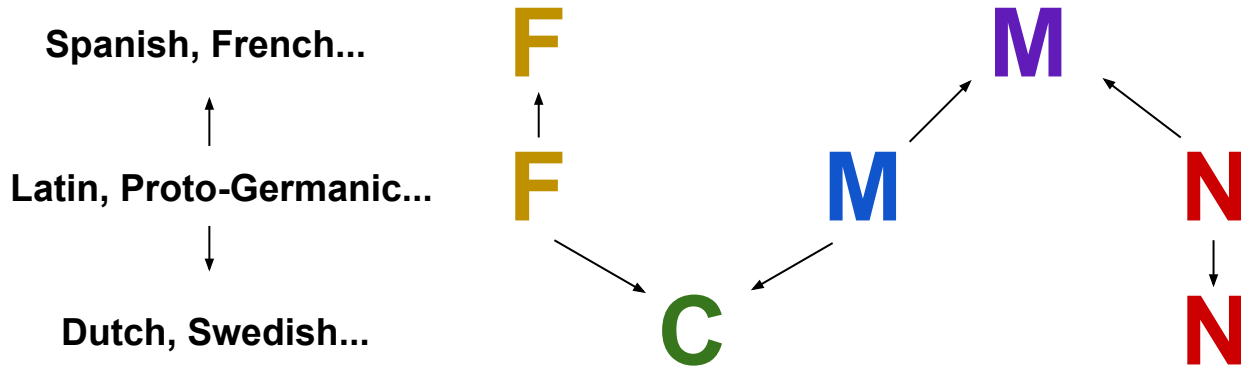
Noun Classes - 2 Genders (revisited)

- **M-F isn't the only possible two gender system.**
- **Many European languages (Dutch, Swedish, etc) make a common-neuter distinction.**
- **This is even farther removed from "natural" gender**

Noun Classes - 2 Genders (revisited)

- **M**-**F** isn't the only possible two gender system.
- **Many European languages (Dutch, Swedish, etc) make a common-neuter distinction.**
- **This is even farther removed from "natural" gender**
- **In Europe, M-F and C-N evolved from M-F-N**

Noun Classes - 2 Genders (revisited)

- **M-F** isn't the only possible two gender system.
- **Many European languages (Dutch, Swedish, etc) make a common-neuter distinction.**
- **This is even farther removed from "natural" gender**
- **In Europe, M-F and C-N evolved from M-F-N**

Spanish, French...

↑

Latin, Proto-Germanic...

↓

Dutch, Swedish...

F

F → M → C ← M

M → C

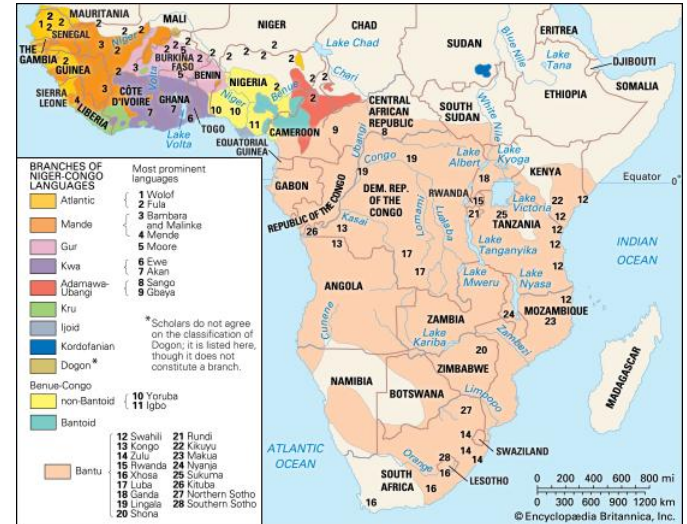N

N

Noun Classes - Lots*!*

- **In some languages, noun classes make no reference to gender at all**

Noun Classes - Lots*!*

- **In some languages, noun classes make no reference to gender at all**
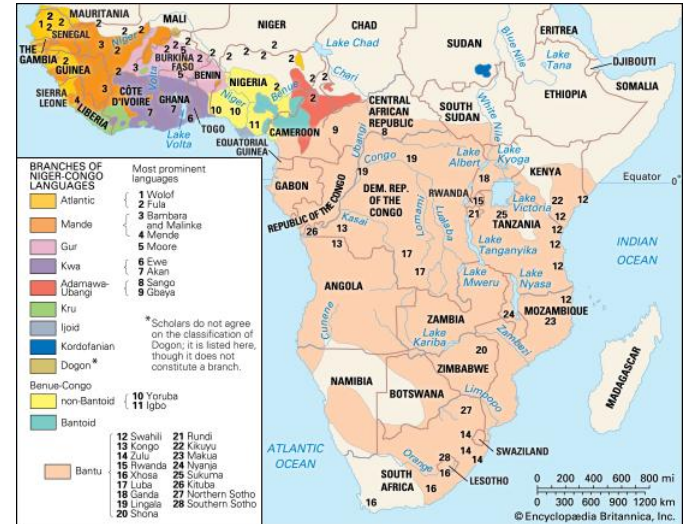- **And some languages have far more than 2 or 3 classes**

Noun Classes - Lots*!*

- **In some languages, noun classes make no reference to gender at all**
- **And some languages have far more than 2 or 3 classes**
- **The Bantu family (sub-Saharan Africa) is famous for this**

Noun Classes - Lots*!*

- **In some languages, noun classes make no reference to gender at all**
- **And some languages have far more than 2 or 3 classes**
- **The Bantu family (sub-Saharan Africa) is famous for this**
- **Most have >10 classes**
- **Indicated by prefixing**

Noun Classes - Shona (Bantu) Example (pt. 1)

Class #	Class Meaning	Example	Singular	Plural
1/2	Person	Shona person	**mu**Shona	**va**Shona
3/4	Plants, stuff	tree	**mu**ti	**mi**ti
5/6	(bad?) stuff, animals	baboon	gudo	**ma**kudo
7/8	Stuff, abstract	corn cob	**chi**bage	**zvi**bage
9/10	stuff, animals	lion	shumba	shumba
11	long stuff, abstract	river	**ru**kova	N/A
12/13	small things, dimn.	insect	**ka**mbuyu	**tu**mbuyu

Noun Classes - Shona (Bantu) Example (pt. 1)

Class #	Class Meaning	Example	Singular	Plural
14	abstract	humanity	hunhu	N/A
15	verbal N, non-fin. V	to speak	kutaura	N/A
16	"near" prepositional	here at home	pamba	N/A
17	"far" prepositional	at home	kumba	N/A
18	"in" prepositional	in the house	mumba	N/A
19/20	*(in some varieties)*			
21	big stuff, augm.	big frog	zura	zura

Extra Credit Quiz