

# **Variability of languages in time and space**

## **Lecture 1: Basic facts and linguistic resources on the world's languages**

Magda Ševčíková

September 29, 2021

- 1 Basic facts
- 2 Variability of languages
- 3 Linguistic resources on the world's languages

- 1 Basic facts
- 2 Variability of languages
- 3 Linguistic resources on the world's languages

- The number of languages in printed and online resources:
  - *Ethnologue* (24th ed., Simons – Fennig 2021): **7,139 languages**
  - *Ethnologue* (23rd ed., Simons – Fennig 2020): 7,117 languages
  - *Ethnologue* (22nd ed., Simons – Fennig 2019): 7,111 languages
  - *Ethnologue* (21st ed., Simons – Fennig 2018): 7,097 languages
  - *Ethnologue* (16th ed., 2009): 6,909 languages
  - *Ethnologue* (7th ed., 1969): 4,493 languages
  - *Encyclopedia Britannica* (1911): app. 1,000 lang.
- not an increase in the number of languages, but rather steadily improving documentation efforts

# First language vs. second language

- **A first language (L1)**

is a language which a person acquires first in her/his life, usually naturally as a child.

- **A second language (L2)**

is a language which a person acquires next to her/his first language.

- **A living language**

has at least one speaker for whom it is a first language.

- e.g. English, Catalan, Japanese, Swahili, French Sign Language

– distinguished from *dead* languages

- a **dead language** is no longer the native language of any community but is still in use, e.g. Latin

– a more subtle insight provided by the

**Expanded Graded Intergenerational Disruption Scale** (EGIDS; Lewis & Simons 2010)

- from *internationally used* languages to *extinct* languages

# Expanded Graded Intergenerational Disruption Scale (EGIDS)

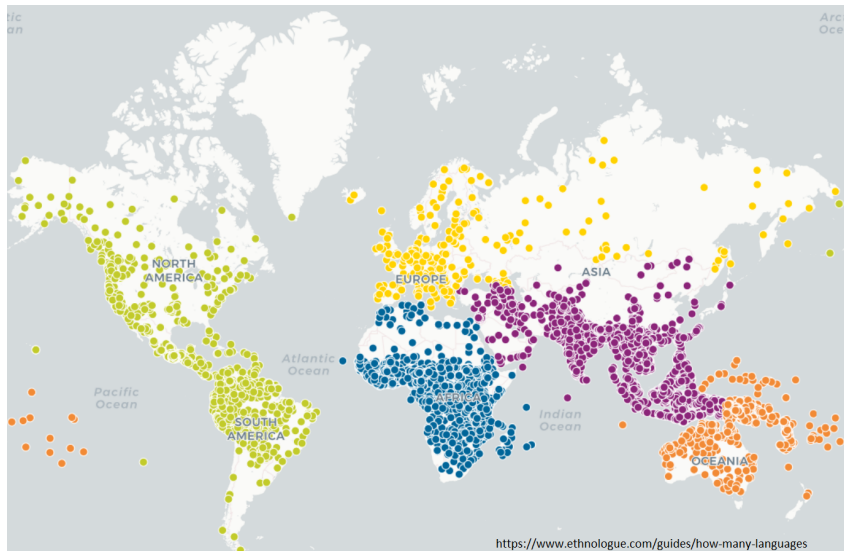
<https://www.ethnologue.com/about/language-status:>

Level	Label	Description
0	International	The language is widely used between nations in trade, knowledge exchange, and international policy.
1	National	The language is used in education, work, mass media, and government at the national level.
2	Provincial	The language is used in education, work, mass media, and government within major administrative subdivisions of a nation.
3	Wider Communication	The language is used in work and mass media without official status to transcend language differences across a region.
4	Educational	The language is in vigorous use, with standardization and literature being sustained through a widespread system of institutionally supported education.
5	Developing	The language is in vigorous use, with literature in a standardized form being used by some though this is not yet widespread or sustainable.
6a	Vigorous	The language is used for face-to-face communication by all generations and the situation is sustainable.
6b	Threatened	The language is used for face-to-face communication within all generations, but it is losing users.
7	Shifting	The child-bearing generation can use the language among themselves, but it is not being transmitted to children.
8a	Moribund	The only remaining active users of the language are members of the grandparent generation and older.
8b	Nearly Extinct	The only remaining users of the language are members of the grandparent generation or older who have little opportunity to use the language.
9	Dormant	The language serves as a reminder of heritage identity for an ethnic community, but no one has more than symbolic proficiency.
10	Extinct	The language is no longer used and no one retains a sense of ethnic identity associated with the language.

# Region of origin of a language

- In maps at *ethnologue.com*, each language is located in its **primary country**
  - each language shown **just once**, e.g.
    - English located in the United Kingdom
    - Esperanto located in Poland

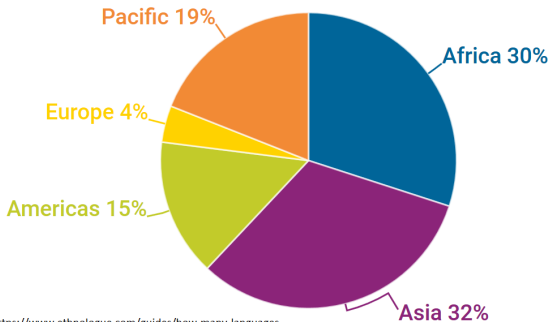
# Regions of origin of languages



# Languages by region of origin

*Ethnologue* (2018):

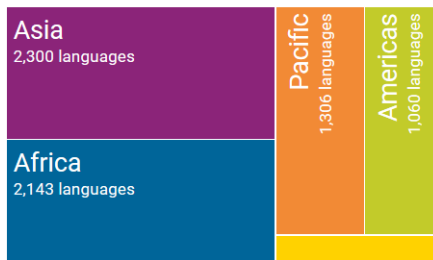
- 2,300 languages (out of 7,097) are from Asia
- 2,143 languages are concentrated in Africa
- 288 languages belong to Europe



<https://www.ethnologue.com/guides/how-many-languages>

# Languages by region and by size (number of L1 speakers)

## Languages



<https://www.ethnologue.com/guides/how-many-languages>

Europe (yellow): 288 languages

## Population



Americas (green): 49M  
Pacific (orange): 7M speakers

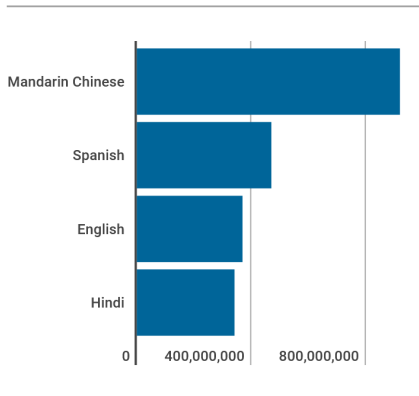
Note: English listed as a language belonging to Europe. Therefore, all L1 speakers of English (e.g. incl. English speakers living in the USA) categorized under Europe.

# Languages by number of L1 speakers

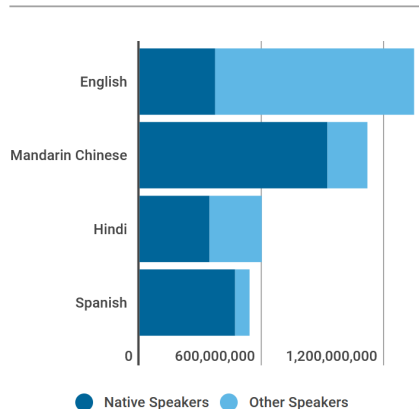
- nearly 7,900,000,000 speakers around the world
  - nearly 90 % of speakers use languages from Asia or Europe as L1
  - 40 % of speakers use one of just 5 languages as L1
    - Chinese, Spanish, English, Hindi, Arabic

# Number of L1 speakers vs. of L2 speakers

*Languages with the most native speakers*



*Languages with the most speakers*



- 1 Basic facts
- 2 Variability of languages
- 3 Linguistic resources on the world's languages

# Variability of languages

- Languages share some features and differ in others
  - due to a common origin
  - due to mutual contact
- Language typology compares
  - phonological features
  - morphology: inflections and word-formation
  - syntax
- Spoken vs written language
- Sociolinguistic aspects
- Stylistics

# Similarities and differences between languages

Languages of the world do share some features, but display differences at the same time.

En.	German	Czech	Slovene	Croatian	French	Spanish	Italian
<i>father</i>	<i>Vater</i>	<i>otec</i>	<i>oče</i>	<i>otac</i>	<i>père</i>	<i>padre</i>	<i>padre</i>

- WHY are they similar?
  - WHY do they differ?
- 
- common origin → genetic classification
  - language contact → areal typology
  - common structural features → language / linguistic typology

# Genetic classification

Some similarities can be explained by the **common origin** of languages.

En.	German	Czech	Slovene	Croatian	French	Spanish	Italian
<b>father</b>	<b>Vater</b>	<b>otec</b>	<b>oče</b>	<b>otac</b>	<b>père</b>	<b>padre</b>	<b>padre</b>

- languages that developed from the same ancestor language and belong to the same **language family** share common features
  - English and German belong to **Germanic** languages
  - Czech, Slovene, and Croatian to **Slavic** languages
  - French, Spanish and Italian to **Romance** languages

But what about similarities between unrelated languages? They can result from **language contact**.

En.	Czech	Basque	Swahili	Maori
<i>pepper</i>	<i>pepř</i>	<i>piperra</i>	<i>pilipili</i>	<i>pepa</i>

“The country of origin of black pepper is south India. The origin of the word can be found in Sanskrit pippali. The merchants from India distributed to other countries not only spice but also its name. In this way languages got in touch, in contact.” (Körvélyessy 2017:80)

Languages that are geographically in contact may influence each other

- based on the regional location of languages
- without respect to their genetic classification
- speakers of the languages interact and influence each other
  - pronunciation
  - grammatical constructions
  - borrowing of words

language	'woman' – ' <b>the</b> woman'	'man' – ' <b>the</b> man'
Bulgarian	жена – жената	мъж – мъжът
Romanian	femeie – fememia	bărbat – bărbatul
Albanian	grua – grua <b>ja</b>	burrë – burri

Körtvélyessy (2017:2):

“**Language typology** is a system or study that divides languages into smaller groups according to similar properties they have. [...] These smaller groups are called **language types**.”

- a **holistic approach** to language typology

“The classification of languages into language types attempts to ‘match’ the complete language system with one language type.”

- a **partial approach** to language typology

“the classification is based on the analysis of a selected language construction and/or phenomenon (not the entire language), for example the size of the consonantal inventory, the presence vs. absence of articles in language, the order of words in a sentence etc.”

- Typological classification is based on the study of structural and functional features at particular levels of the language system.
- Typological approach to phonology:
  - How many vowels and consonants are used in the language?
  - Which features do they have?
  - How are syllable formed in the language?
  - Which prosodic features are employed in the language?

- Are grammatical categories conveyed by morphemes, auxiliaries, and/or by word order?
  - How many morphological cases are available in the language?
- In holistic approaches to morphological typology, several language types discerned:
- ① analytic – isolating
  - ② synthetic – agglutinating
  - ③ synthetic – inflectional
  - ④ polysynthetic
  - ⑤ introflective

# Language typology of word formation

- Which word-formation processes are available in a language to coin new words?
- How frequent are particular word-formation processes in a language?
- How are things/abstract phenomena/actions named in a language?

- How are sentence elements ordered in a linear sequence (sentence)?
  - focus on three main elements: *subject, verb, object*
  - the order of the elements analysed in simple indicative sentences

## A language universal

is a feature that all languages have, or a statement that holds for all languages, or *less strictly* (more commonly), a feature or statement that holds for a majority of languages.

- morphological and syntactic universals listed in the *Universals Archive*  
<http://typo.uni-konstanz.de/archive/intro>

# Spoken vs written communication

- What syntactic structures can we expect in spoken communication?
- What is the density of information in spoken in written text?
- How is written text encoded (writing systems)?

- How do we choose from varieties of a language?
- How is the usage of languages / varieties connected with social variables (age, nationality, education...)?
- What is the speakers' relation to the language norm?

# Stylistics: possibility of choice

- What are the features of different language registers (e.g., in administration, media, informal talk)?
- What are the features of text genres (e.g., formal letter, joke, detective story)?
- What is the usual setting of information in different text genres?

- 1 Basic facts
- 2 Variability of languages
- 3 Linguistic resources on the world's languages

# Online resources on the world's languages

- *Ethnologue: Languages of the World*
  - <http://www.ethnologue.com>
- **WALS – The World Atlas of Language Structures**
  - <http://wals.info>
- *APiCS – Atlas of Pidgin and Creole Language Structures*
  - <http://apics-online.info>
- *WOLD – The World Loanword Database*
  - <http://wold.clld.org>
- for other resources see e.g.
  - <https://linguistic-typology.org/databases/>

- database of phonological, grammatical and lexical properties of languages
  - obtained from reference grammars and other descriptive material
- 55 authors
  - Greville G. Corbett, Martin Haspelmath, Bernard Comrie, Matthew S. Dryer et al.
- editions
  - 1st version 2005
    - book with CD-ROM, Oxford University Press
  - 1st online version (*WALS Online*) 2008
    - a separate publication of the Max Planck Institute for Evolutionary Anthropology
  - current version from 2013
    - <http://wals.info>

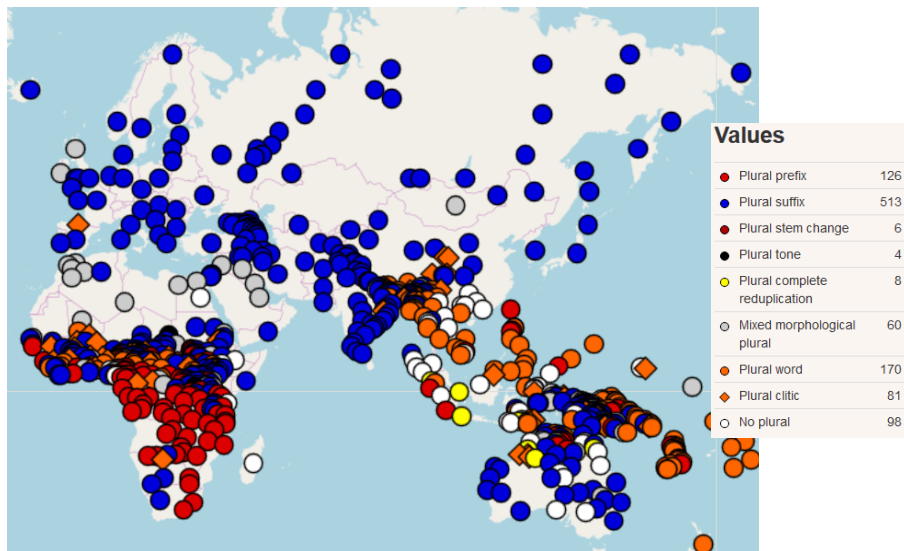
# Linguistic features in *WALS*

- 144 features in total
  - structural properties of language that describe “one aspect of linguistic diversity”
  - concise linguistic description of each feature
  - 2 to 28 values of the feature
  - assigned to one of 11 areas:

phonology	nominal syntax	lexicon
morphology	simple clauses	sign languages
nominal categories	complex sentences	other
verbal categories	word order	

- distribution of the feature values on the map

# WALS feature 33A: Coding of Nominal Plurality



# Description of languages in WALS



THE WORLD ATLAS  
OF LANGUAGE STRUCTURES  
ONLINE



Home Features Chapters Languages References Authors

Family: Indo-European / Genus: Germanic

Glottocode: dutc1256

ISO 639-3: nld

## Language Dutch

WALS code: dut

Showing 1 to 89 of 89 entries

Fid <sup>a</sup>	Value	Feature	Source	Area
	<input type="text" value="Search"/>	<input type="text" value="Search"/>		
14A	No fixed stress	Fixed Stress Locations	Kager 1989; Trommelen and Zonneveld 1999a; van der Hulst 1984	Phonology
15A	Right-oriented: One of the last three	Weight-Sensitive Stress	Kager 1989; Trommelen and Zonneveld 1999a; van der Hulst 1984	Phonology
16A	Coda consonant	Weight Factors in Weight-Sensitive Stress Systems	Kager 1989; Trommelen and Zonneveld 1999a; van der Hulst 1984	Phonology
17A	Trochaic	Rhythm Types	Kager 1989; Trommelen and Zonneveld 1999a; van der Hulst 1984	Phonology
26A	Strongly suffixing	Prefixing vs. Suffixing in	Koolhoven 1968:	Morphology



Coordinates WGS84 52°30'N, 6°E  
52.50, 6.00

Spoken in: Netherlands

### Alternative names

Ruhlen: Dutch

Routledge: Netherlandic

Ethnologue: Dutch

### Sources

## References

- Comrie, B., Matthews, S. & Polinsky, M.: *The Atlas of Languages: The Origin and Development of Languages Throughout the World*. 1997.
- Comrie, B. (ed.): *The World's Major Languages*. Oxford: Oxford University Press, 1987.
- Dryer, M. S. & Haspelmath, M. (eds., 2013): *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online: <http://wals.info>
- Haspelmath, M. et al. (eds.; 2001): *Language Typology and Language Universals*. Mouton de Gruyter.
- Körtvélyessy, L. (2017): *Essentials of Language Typology*. Košice: UPJŠ. Available online: <https://unibook.upjs.sk/sk/anglistika-a-amerikanistika/365-essentials-of-language-typology>
- Lewis, M. P. & Simons, G. F. (2010): Assessing endangerment: Expanding Fishman's GIDS. *Revue Roumaine de Linguistique*, 55, pp. 103–120.
- Matthews, P. H. (2007): *The Concise Oxford Dictionary of Linguistics*. Oxford: Oxford University Press.
- Simons, G. F. & Fennig, C. D. (eds., 2018): *Ethnologue: Languages of the World*, Twenty-first edition. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com>

- other online resources:

- *APiCS – Atlas of Pidgin and Creole Language Structures*. Available online: <http://apics-online.info>
- *WOLD – The World Loanword Database*. Available online: <http://wold.clld.org>
- <https://www.linguisticsociety.org>
- <http://www.linguistic-typology.org>
- <https://www.omniglot.com>
- [www.uis.unesco.org](http://www.uis.unesco.org)