

Variability of Languages in Time and Space

Intro Phonological Typology Consonant and Vowel inventories

- Motivation for Linguistic Typology
- Phonetics and Phonology
- Consonant and Vowel Inventories
- Segmental Processes
- Linguistic quiz to the end

Magda Sevcíková & Anna Nedoluzhko & Zdeněk Žabokrtský

Formalities

Variability of languages in time and space (NPFL150)

Teachers: Magda Ševčíková, Anja Nedoluzhko, Zdeněk Žabokrtský

Fall term 2025/2026: Thursday 10:40-12:10 in S11

Course completion requirements: 3 homework assignments and a final test

- **3 homework assignments** (Zdeněk Žabokrtský)

For each assignment, you will get points, up to a given maximum (specified with each assignment).

If your submission is especially good, you can get extra points (up to +10% of the maximum).

All assignments will have a fixed deadline (usually in 10 days).

If you submit the assignment after the deadline, you will get:

up to 50% of the maximum points if it is less than 2 weeks after the deadline;

0 points if it is more than 2 weeks after the deadline.

To be allowed to take the test (which is required to pass the course), you need to get at least 50% of the total points from the assignments.

- **Final test**

There will be a written test at the end of the semester.

To pass the course, you need to get at least 50% of the total points from the test.

Living Languages Around The World

The number of languages in printed and online resources:

Ethnologue (28th ed., 2025): 7,159 languages

Ethnologue (27th ed., 2024): 7,164 languages

Ethnologue (26th ed., 2023): 7,168 languages

Ethnologue (25th ed., 2022): 7,151 languages

Ethnologue (24th ed., 2021): 7,139 languages

Ethnologue (23rd ed., 2020): 7,117 languages

Ethnologue (22nd ed., 2019): 7,111 languages

Ethnologue (21st ed., 2018): 7,097 languages

...

Ethnologue (16th ed., 2009): 6,909 languages

...

Ethnologue (7th ed., 1969): 4,493 languages

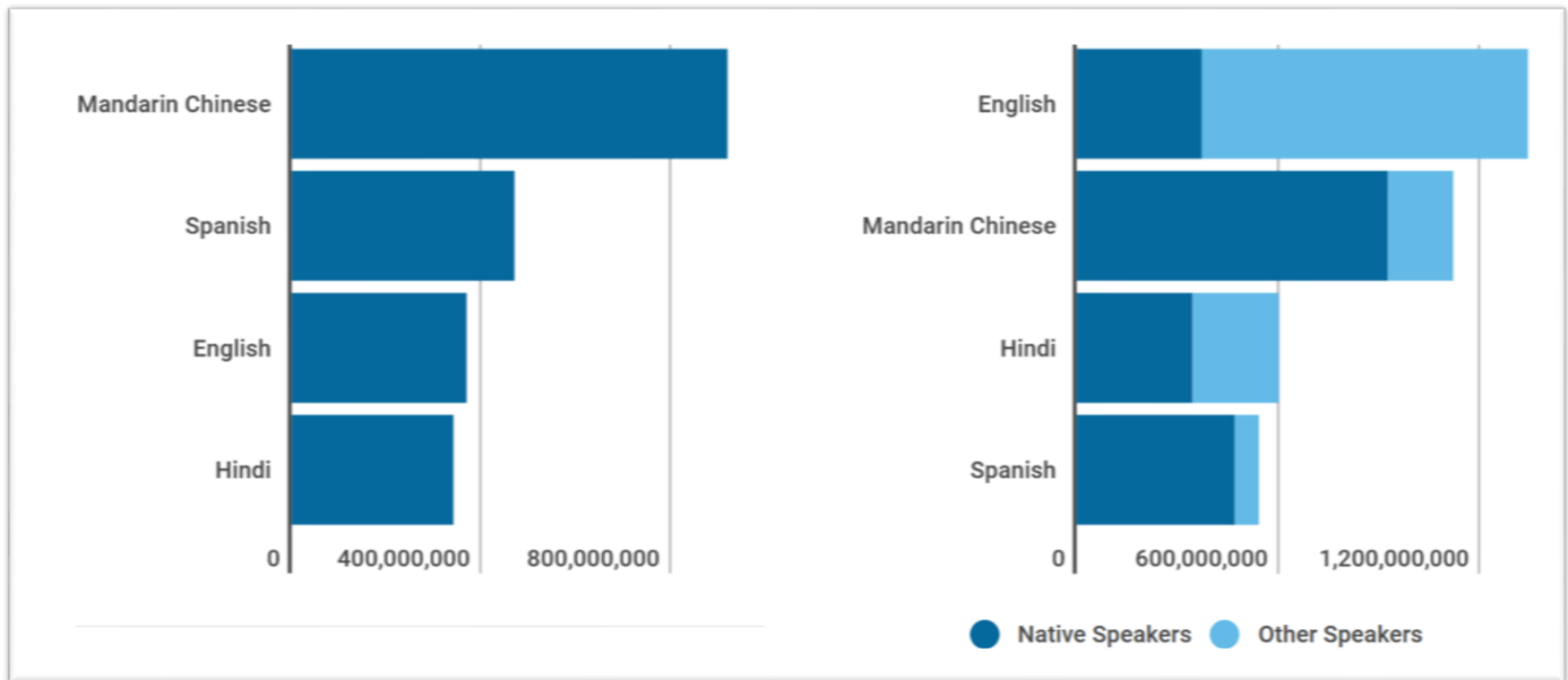
...

Encyclopedia Britannica (1911): app. 1,000 lang.

- no increase in the number of languages, rather improving documentation efforts
- aligns to the ISO 639-3 standard

Number of L1 speakers vs. L2 speakers

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



Variability of languages

- Languages of the world share some features, but they also differ in many ways
 - WHY are they similar? WHY do they differ?

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>

- due to a common origin → genealogical (genetic) classification
- due to language contact → areal linguistics – due to other factors
- due to other factors

Genealogical classification

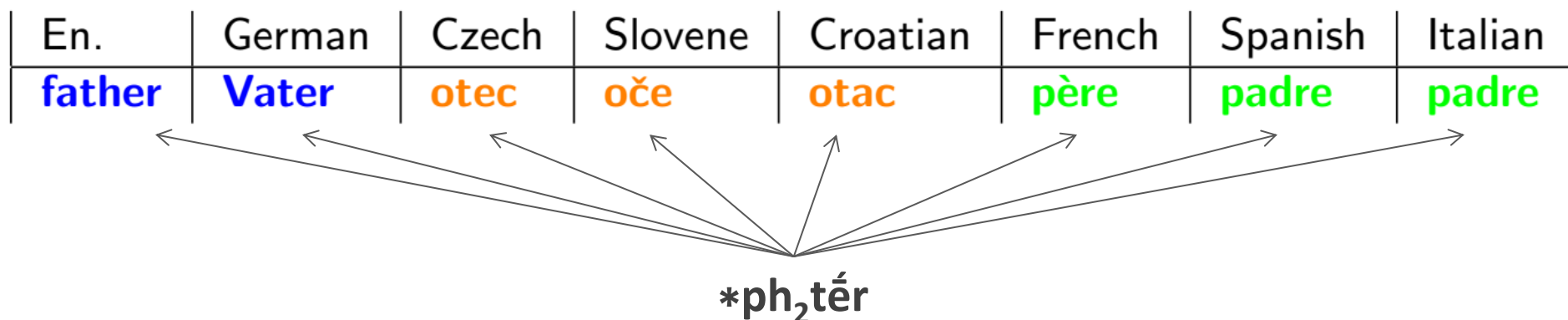
Some similarities can be explained by the **common origin** of languages.

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

- 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

Genealogical classification

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



- Languages that display systematic similarities are assumed to have descended from *a common source language* = **ancestor language / proto-language**.
 - regular correspondences between languages studied by comparative methods
 - proto-languages not attested, they are **reconstructed** from available documents of the descendent languages

Language Families

Language family is a group of languages that have developed from the same proto-language.

language family	languages	million of speakers
Niger-Congo	1,527	511
Austronesian	1,224	324
Trans-New Guinea	479	3.5
Sino-Tibetan	453	1,370
Indo-European	444	3,090
Afro-Asiatic	366	496

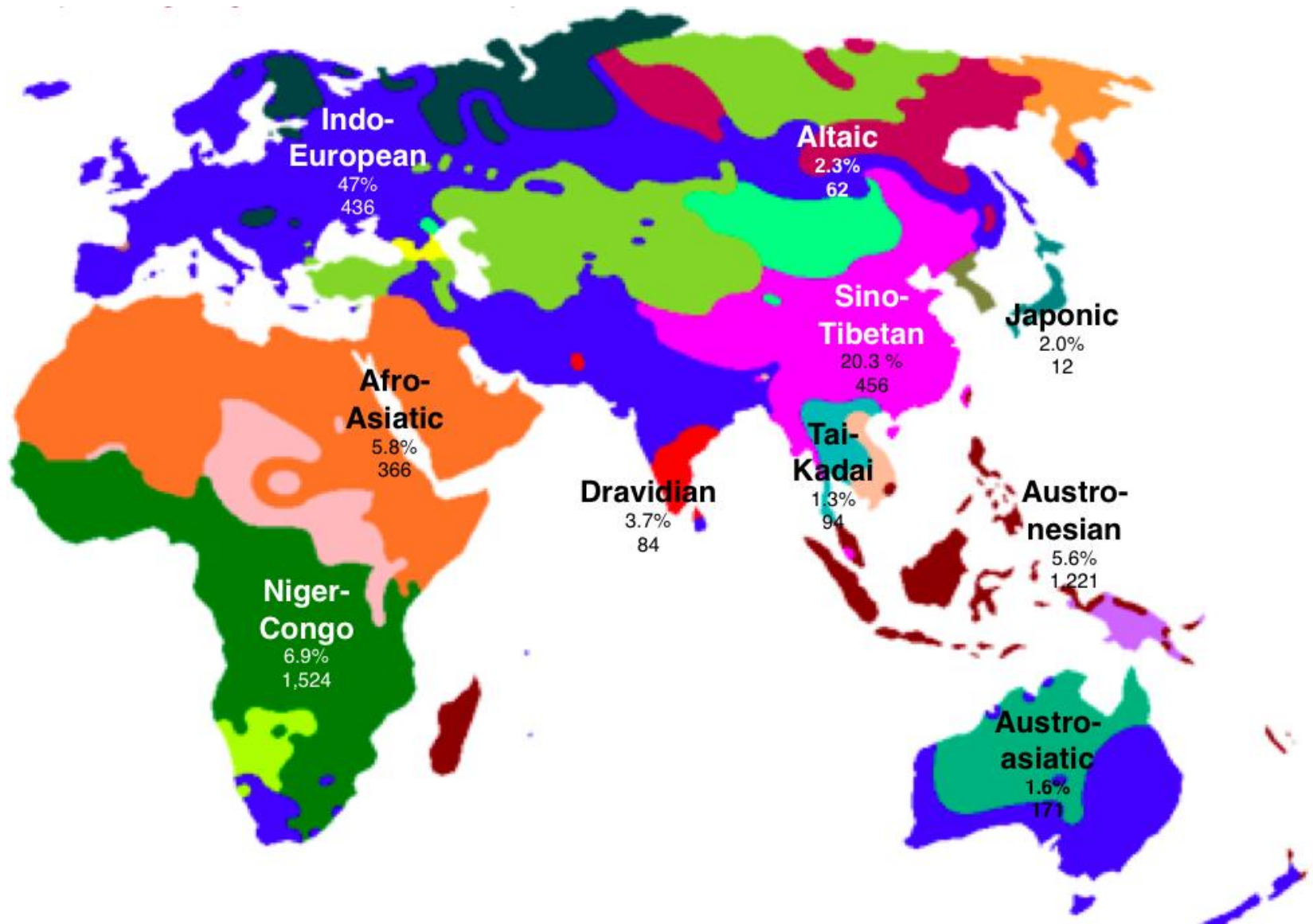
- plus another more than 150 language families in *Ethnologue* 2025:

<https://www.ethnologue.com/browse/families/>

Language isolates are natural languages that cannot be assigned to any language family; e.g. Basque (= Euskera; spoken in Spain), Ainu (Japan).

- sometimes hard to prove (loss of traces/time depth, intense language contact, little data)
 - Japanese, Korean (isolated vs. Altaic), Georgian (Kartvelian, Nostratic hypothesis), Basque (isolated vs. Dené–Caucasian hypothesis)

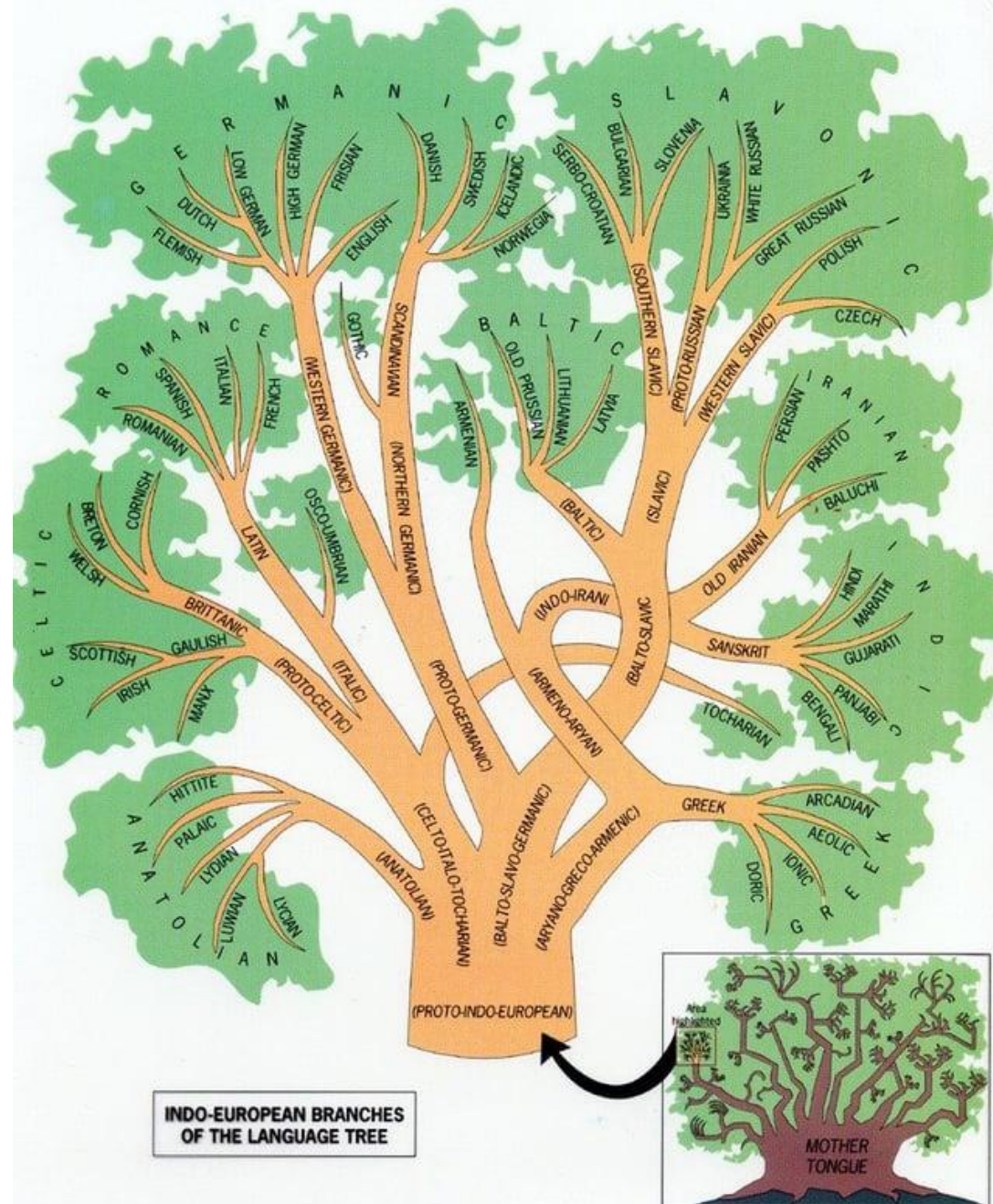
Language Families



Indo-European language family tree

proto-language: Proto-Indo-European (PIE)

- reconstructed on the basis of Sanskrit, Ancient Greek, Latin, and others
- assumed existence: 4,500–2,500 B.C.E.



Limits of Language Tree

- There are significant challenges in language classification
 - 20 Slavic languages in *Ethnologue* vs. 17 Slavic languages in *WALS*

<i>Ethnologue</i>	<i>WALS</i>	
Belarusian	Belorussian	- Croatian
Russian	Russian	- Montenegrin
Ukrainian	Ukrainian	- Church Slavonic
Bulgarian	Bulgarian	- Slavomolisano
Macedonian	Macedonian	- Silesian
Bosnian	Bosnian	- Rusyn
Serbian	Serbian-Croatian	
Slovene	Slovene	
Czech	Czech	<i>WALS</i> only:
Slovak	Slovak	- Sorbian
Kashubian	Kashubian	- Polabian
Polish	Polish	- Slovincian
Lower Sorbian	Lower Sorbian	
Upper Sorbian	Upper Sorbian	

Language Contact

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)

Areal Linguistics

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
 - morphosyntax and syntax

language	'woman' – 'the woman'	'man' – 'the man'
Bulgarian	жена – жената	мъж – мъжът
Romanian	femeie – fememia	bărbat – bărbatul
Albanian	grua – gruaja	burrë – burri

Linguistic Typology

Finnish

talo = house

talo-ssa = in the house

talo-ssa-ni = in my house

talo-ssa-ni-kin = even in my house

Japanese

hana = flower

hana-ga = flower + subject marker

hana-ga-kirei-da = the flower is beautiful

hana-ga-kirei-da-tta = the flower was beautiful

Turkish

ev = house

ev-ler = houses

ev-ler-imiz = our houses

ev-ler-imiz-den = from our houses

Linguistic Typology

Typological classification is based on the study of structural and functional features at particular levels of the language system.

- Phonological features:
 - 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?
- Morphological features:
 - Are grammatical categories conveyed by morphemes, auxiliaries, and/or by word order?
 - How many morphological cases are available in the language?
- Syntactic features:
 - 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

Linguistic Typology - Motivation

Linguistic analysis and comparison of language features on a large number of languages

- gives awareness of what is
 - possible
 - possible and frequent
 - almost or fully impossible
- lets us combine features, find correlations, implications
 - few consonants $\approx$ simple syllables
 - nasal vowels $\rightarrow$ oral vowels
 - *Example:* Many languages have only oral vowels (Spanish, German) and many languages have both oral and nasal vowels (French, Portuguese), but no language has only nasal vowels: the existence of nasal vowels implies the existence of oral vowels.
- helps us understand phonological data in a given language
 - *I've seen this before... Typically it goes like this...*

Linguistic Typology - Motivation

- Informs what is common/uncommon

Odden 2013: „It is very difficult to refuse a claim of the form „X is more common than Y,” except if a very detailed numerical study is undertaken.“

- ‘basic’ SVO word order
- anaphoricity of the definite article
- unstressed vowel reduction, length and stress correlation
- labialized u and ü
- Markedness
 - a comparative concept
 - marked = something atypical, variants, special cases
 - *Example:* In phonology, not all segments have equal status