

Der Schatz der Sprachen

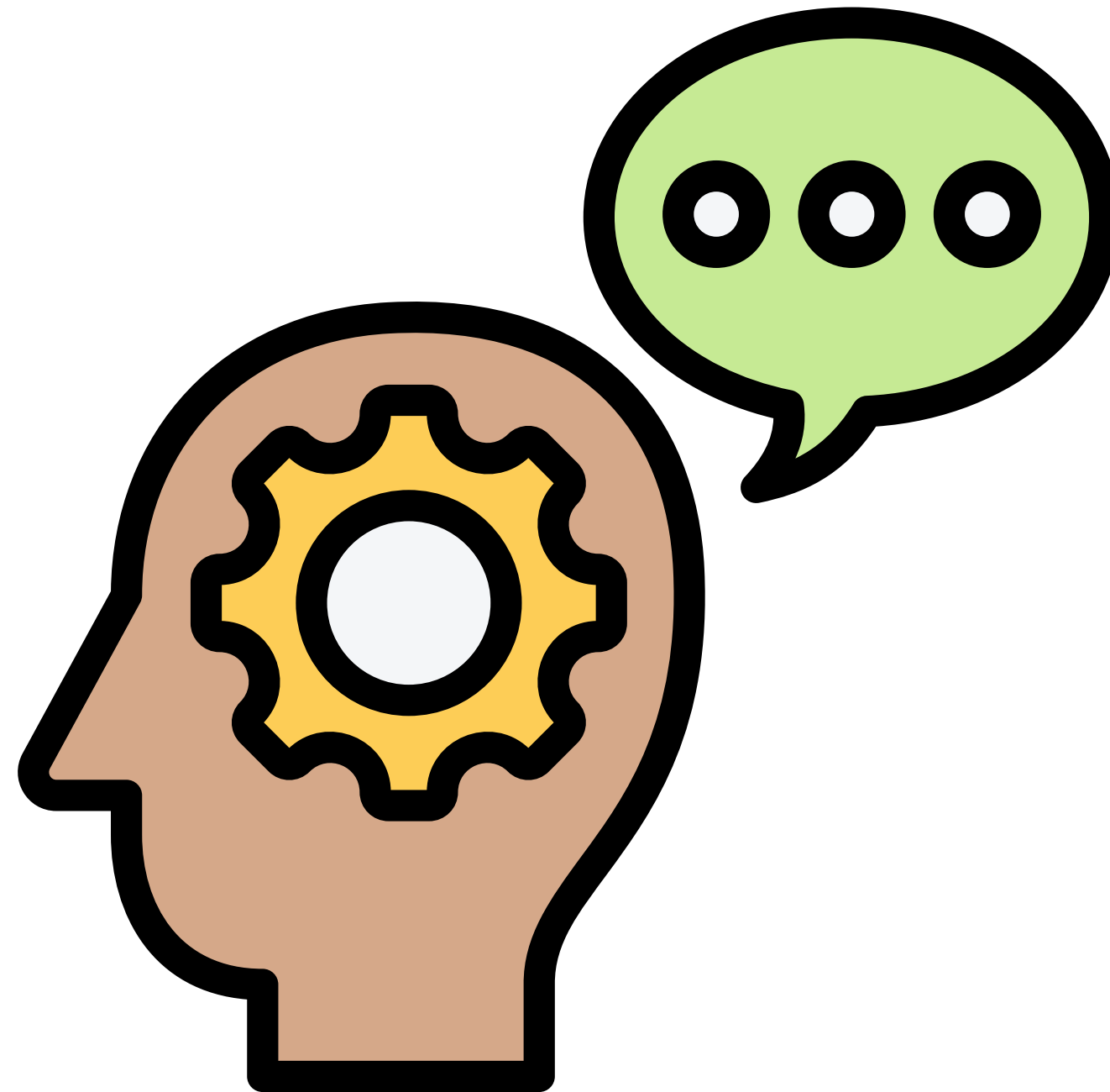
Warum wir Wortschätze in
Datenbanken sammeln

Annika Tjuka

Abteilung für Sprach- und Kulturevolution

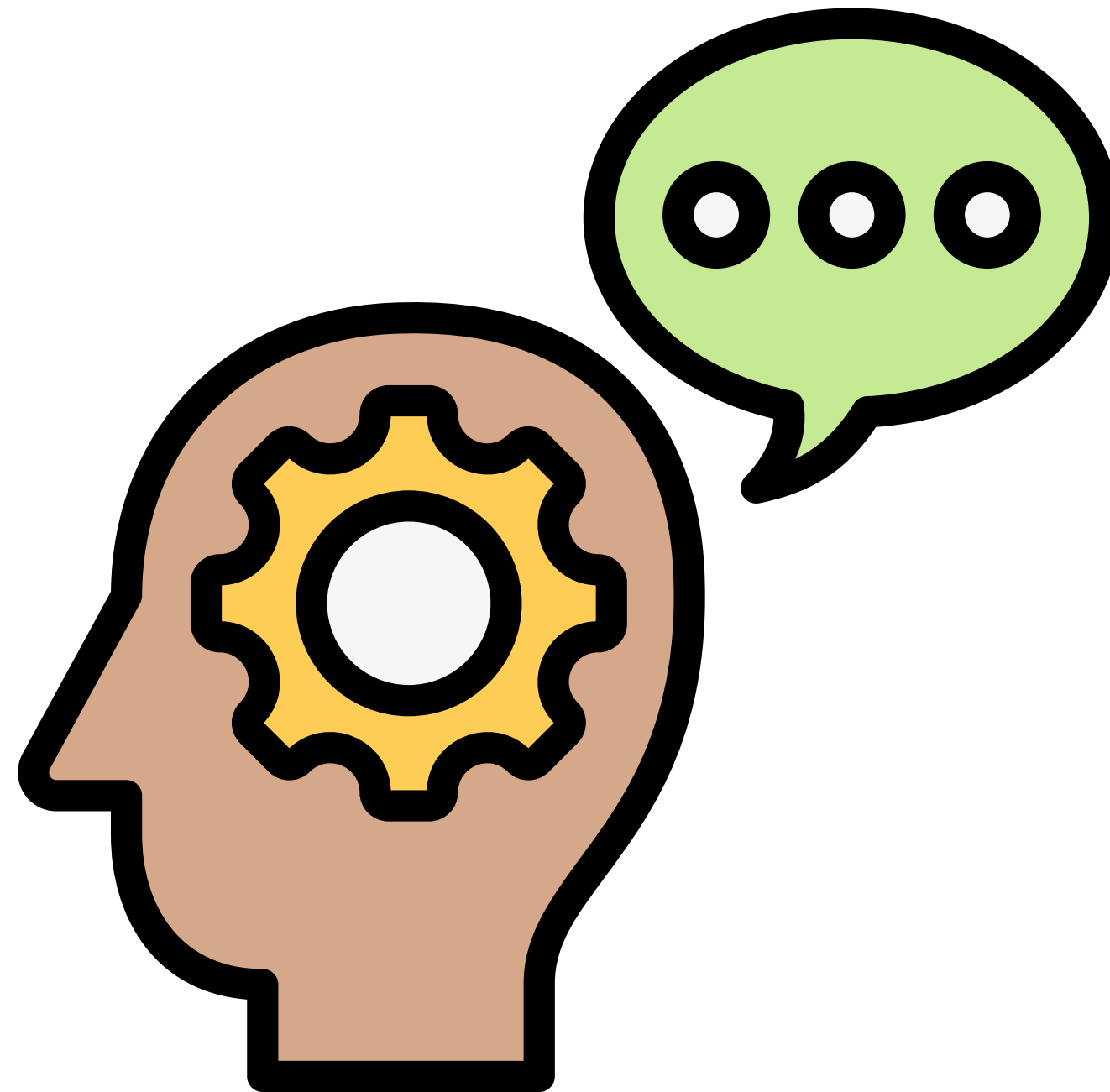


Sprache als System



Struktur

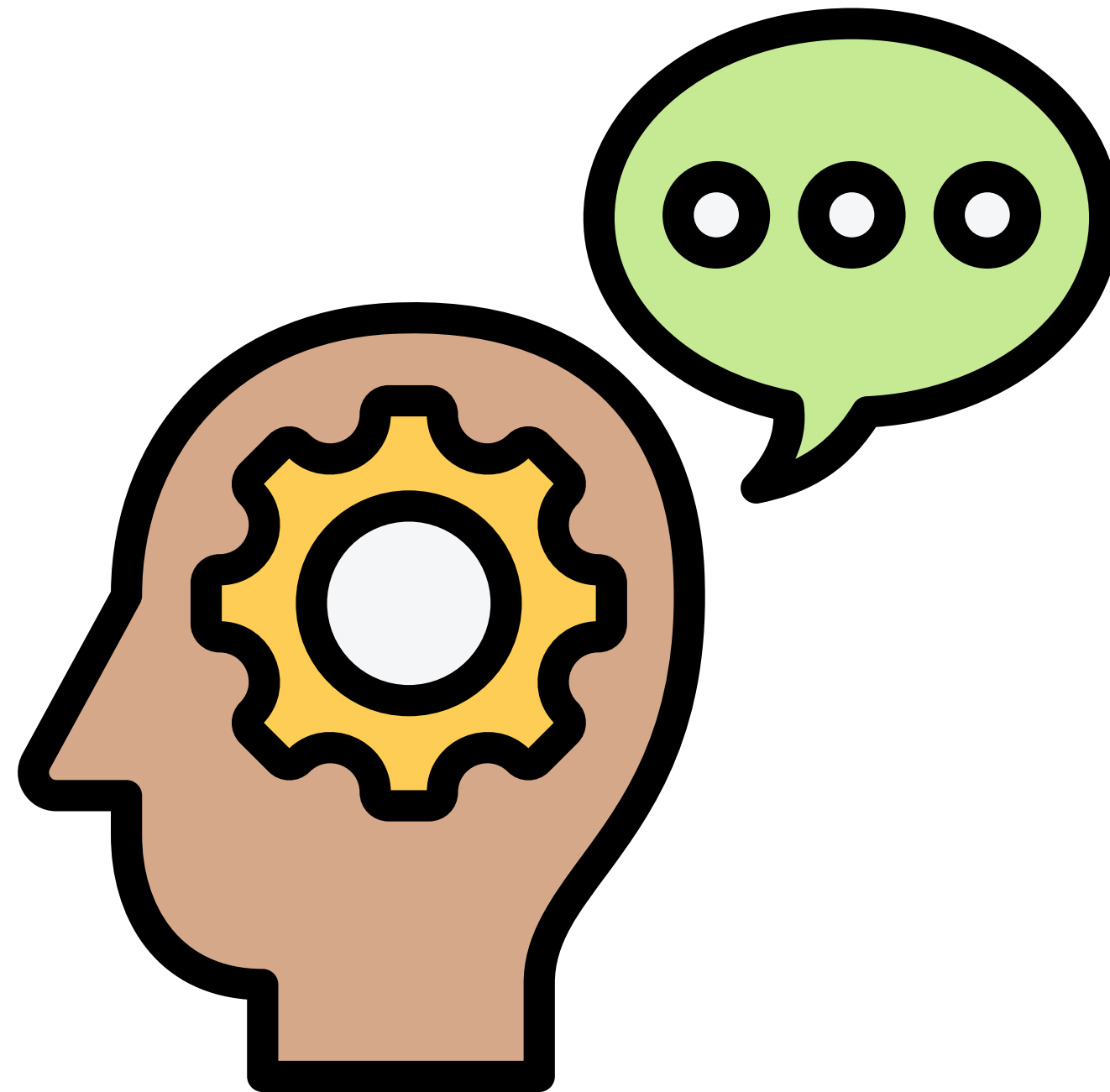
Sprache als System



Struktur

Funktion

Sprache als System

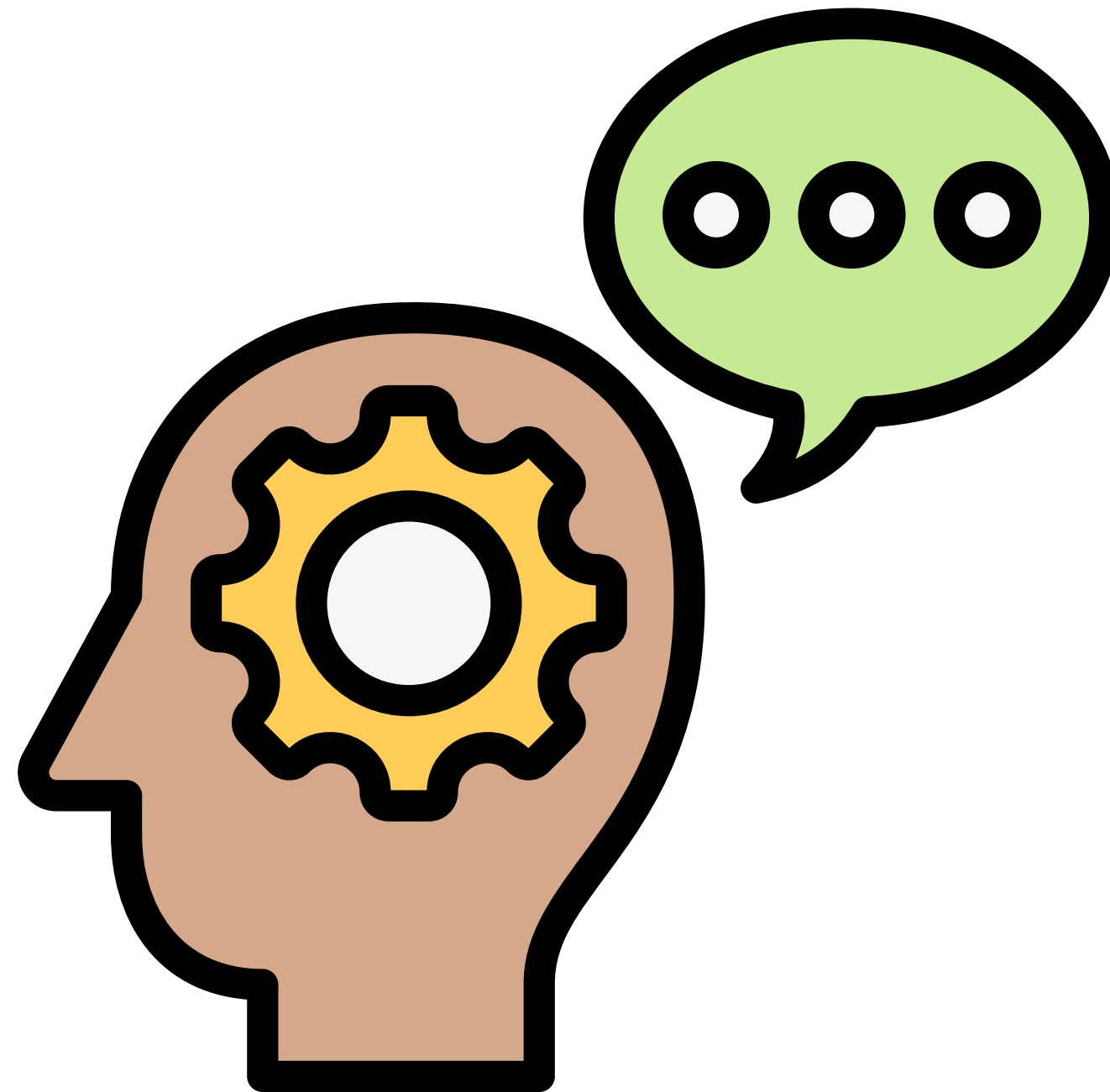


Struktur

Funktion

Entwicklung

Sprache als System



Struktur

Funktion

Entwicklung

Gebrauch

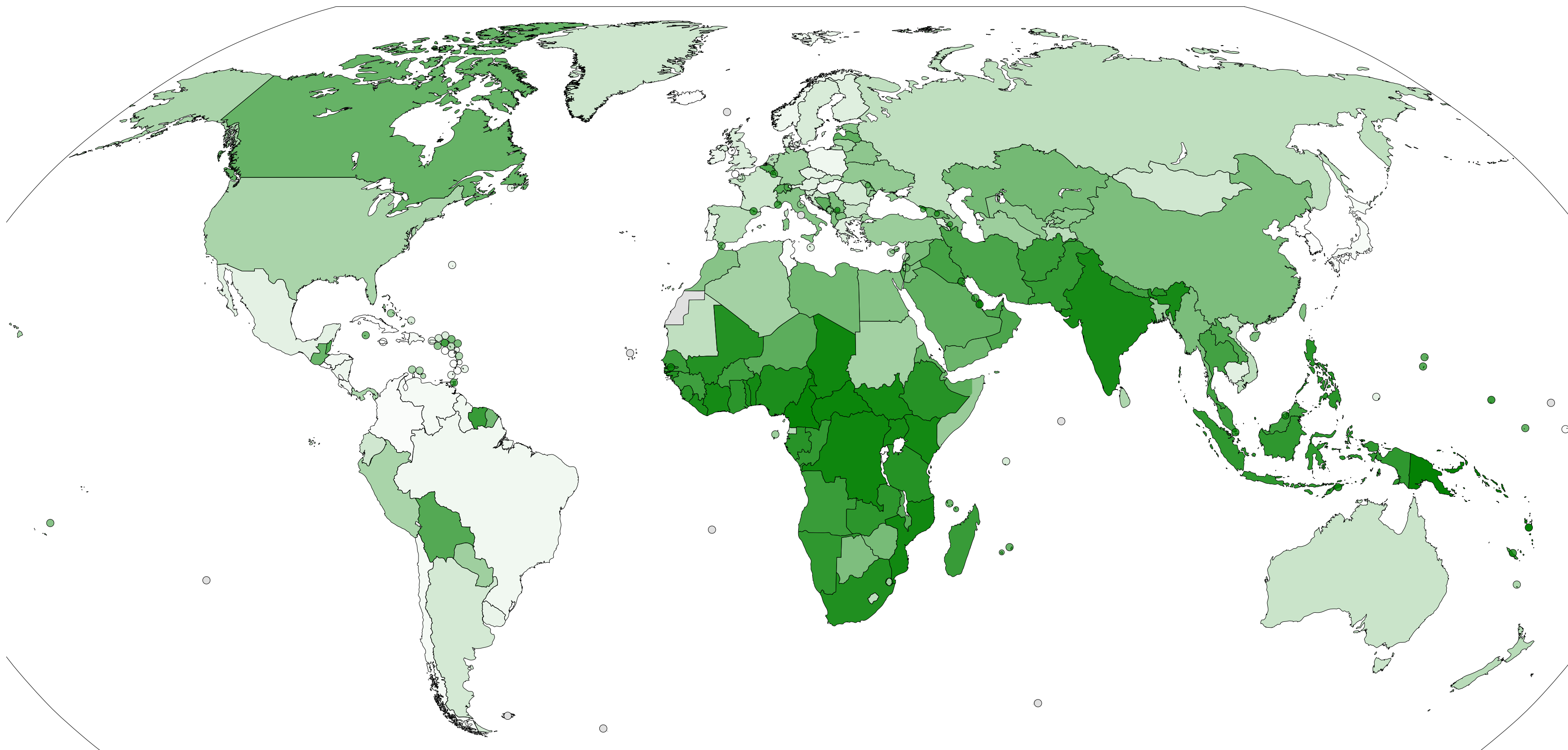
Wie viele Sprachen gibt es in der Welt?

Wie viele Sprachen gibt es in der Welt?

ca. 6.500

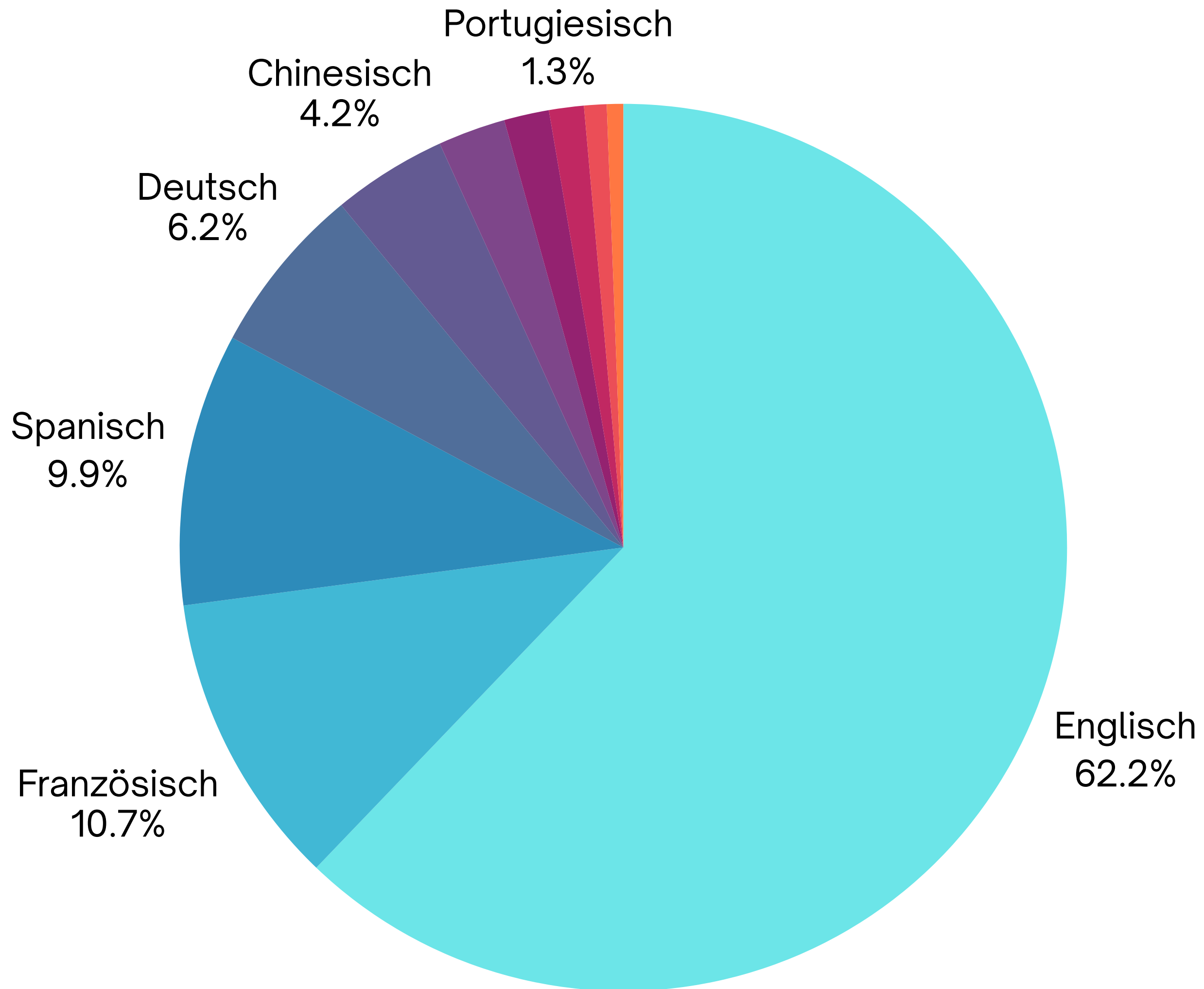


Sprachvielfalt



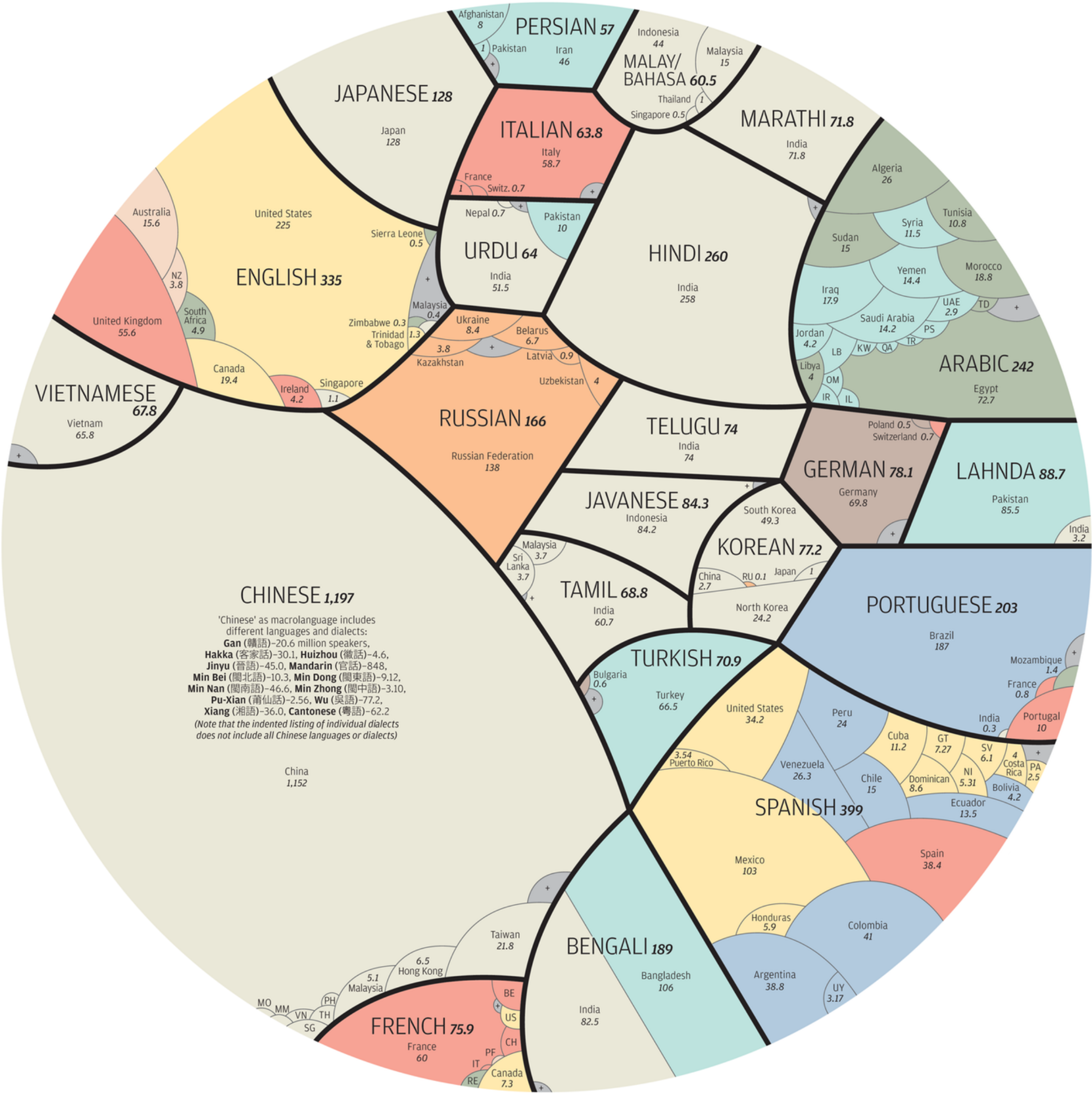
Welche Sprache ist am besten erforscht?

Studien zu Sprachen

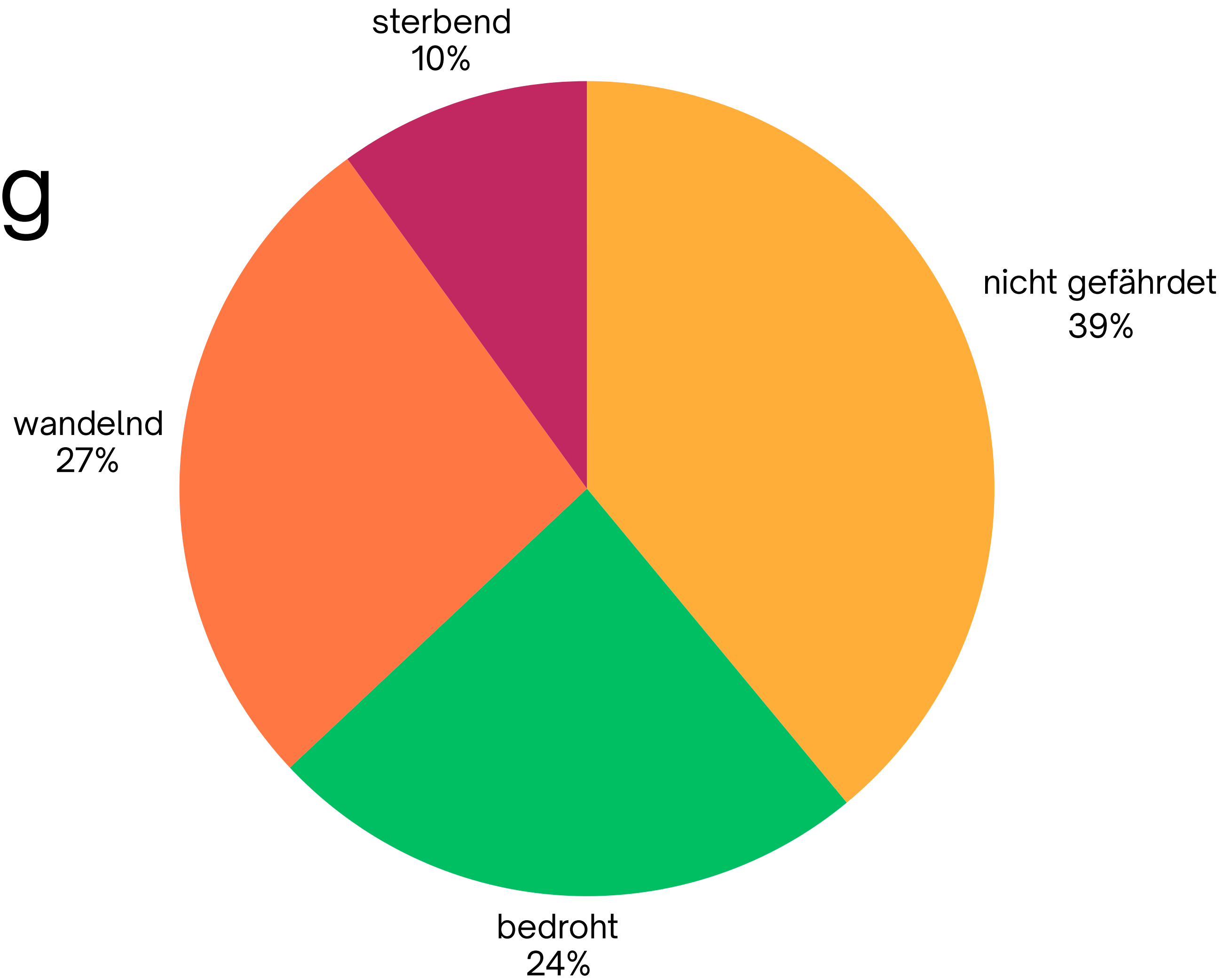


Ist Englisch die Weltsprache?

Anzahl von Sprecher:innen pro Sprache



Gefährdung



Was macht die Abteilung für
Sprach- und Kulturrevolution?

Sprachdokumentation



Vanuatu

Mary Walworth



Ghana

Enock Appiah Tieku



India

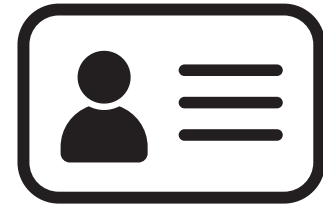
Kumari Mamta

Ergebnis:

Wörterbücher und Grammatiken



Daten

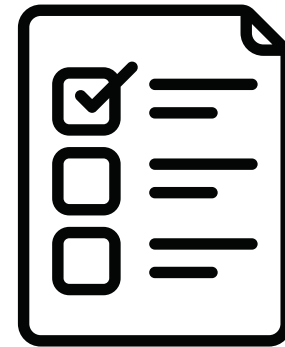


Sprachname

Daten



Sprachname

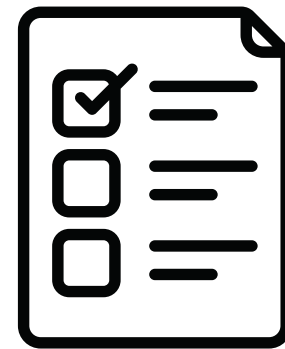


Wörter

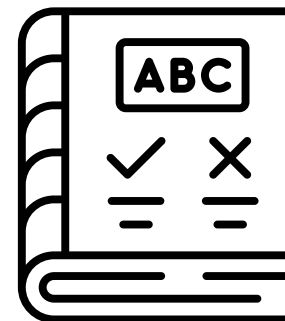
Daten



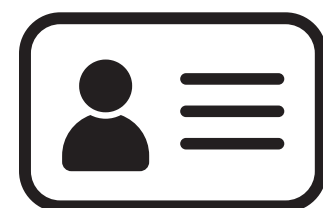
Sprachname



Wörter



Grammatik



Glottolog

Glottolog

Languages

Families

Language Search

References

Reference Search

GlottoScope

About

Glottolog 5.2

Welcome to Glottolog 5.2

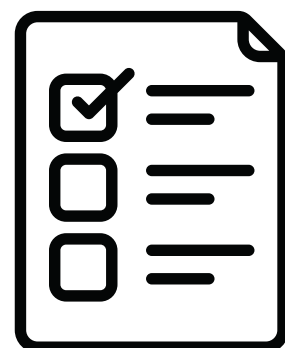
Comprehensive reference information for the world's languages, especially the lesser known languages.

Information about the different languages, dialects, and families of the world ('languoids') is available in the [Languages](#) and [Families](#) sections. The [References](#) section contains bibliographical information. You can query the bibliographical database by filtering the table view or using [a complex query](#) involving genealogical affiliation, document type, and macro-area.

Catalogue of languages and families

Glottolog provides a [comprehensive catalogue](#) of the world's languages, language families and dialects. It assigns a unique and stable identifier (the Glottocode) to (in principle) all languoids, i.e. all families, languages, and dialects. Any variety that a linguist works on should eventually get its own entry. The languoids are organized via a genealogical classification (the Glottolog tree) that is based on available historical-comparative research (see also the [Languoids information](#) section).





Lexibank



lexibank

Home

Datasets

Concepts

Varieties

Words

Legal

Download

Contact

Welcome to Lexibank

Lexibank is different things:

- A framework to curate lexical data. (See <https://github.com/lexibank/>.)
- A [Zenodo community](#), collecting released Lexibank datasets.
- This cld web application, serving an annotated aggregation of selected data from [75 Lexibank datasets](#).

How to cite Lexibank

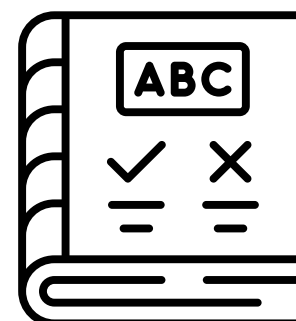
Please cite the Lexibank paper

List, JM., Forkel, R., Greenhill, S.J. et al. Lexibank, a public repository of standardized wordlists with computed phonological and lexical features. *Sci Data* 9, 316 (2022). DOI: [10.1038/s41597-022-01432-0](https://doi.org/10.1038/s41597-022-01432-0)

as well as the underlying dataset

Johann-Mattis List, Robert Forkel, Simon J. Greenhill, Christoph Rzymiski, Johannes Englisch, & Russell D. Gray. (2023). Lexibank Analysed [Data set]. In *Scientific Data* (v1.0, Vol. 9, Number 316, pp. 1–31). Zenodo. DOI: [10.5281/zenodo.7836668](https://doi.org/10.5281/zenodo.7836668)





Grambank

Welcome to Grambank

Grambank was constructed in an international collaboration between the Max Planck institutes in Leipzig and Nijmegen, the Australian National University, the University of Auckland, Harvard University, Yale University, the University of Turku, Kiel University, Uppsala University, SOAS, the Endangered Languages Documentation Programme, and over a hundred scholars from around the world. Grambank is designed to be used to investigate the global distribution of features, language universals, functional dependencies, language prehistory and interactions between language, cognition, culture and environment. The Grambank database currently covers 2,467 language varieties, capturing a wide range of grammatical phenomena in 195 features, from word order to verbal tense, nominal plurals, and many other well-studied comparative linguistic variables. Grambank's coverage spans 215 different language families and 101 isolates from all inhabited continents. The aim is for Grambank to ultimately cover all languages for which a grammar or sketch grammar exists. Grambank is part of Glottobank, a research consortium that involves work on complementary databases of lexical data, paradigms, numerals and sound patterns in the world's languages. Grambank can be used in concert with other databases, such as those in Glottobank and D-PLACE, to deepen our understanding of our history and communicative capabilities.

How to cite Grambank

Please see instructions here: <https://github.com/grambank/grambank/wiki/Citing-grambank>

Data availability

The current release version of the Grambank data can be downloaded from <https://doi.org/10.5281/zenodo.7740139>

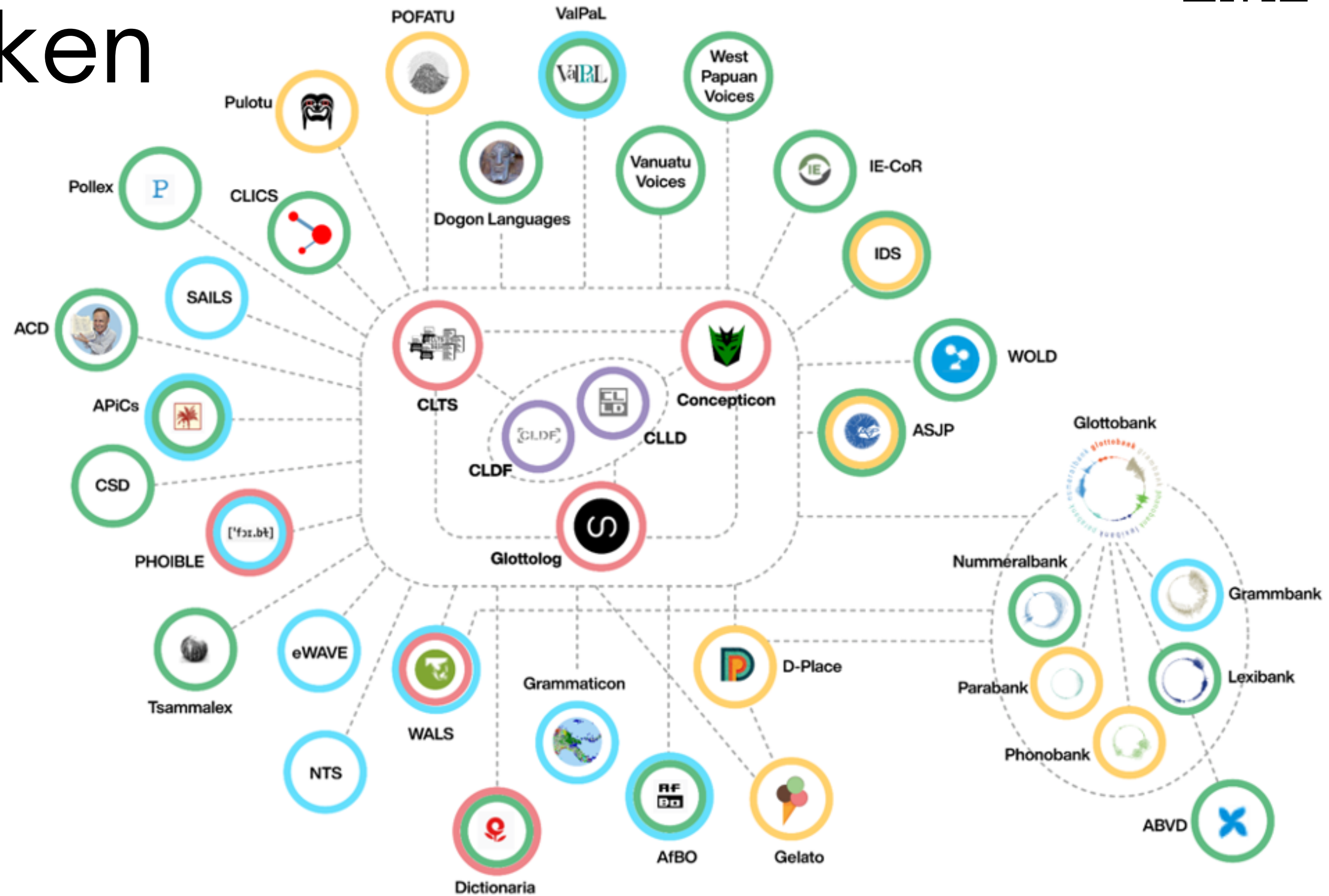
Grambank is a part of the Cross-Linguistic Linked Data-project (CLLD). As such, there will continuously be new versions released. As with all CLLD-databases, it is important that you note down what version you have used in any analysis of the dataset.

Funding

Grambank is a publication of the Department of Linguistic and Cultural Evolution at the Max Planck Institute for Evolutionary Anthropology, Leipzig. Additional funding was provided by the Max Planck Institute for Psycholinguistics in Nijmegen and a Royal Society of New Zealand Marsden grant (UOA1308) to Quentin Atkinson and Russell Gray, and an Australian Research Council Centre of Excellence Grant (CE140100041) for the ARC Centre of Excellence for the Dynamics of Language. The data furnished by the Hunter-Gatherer Language Database was supported by National Science Foundation grant HSD-0902114 'Dynamics of Hunter Gatherer Language Change' PIs Claire Bowern, Patience Epps, Jane Hill, and Keith



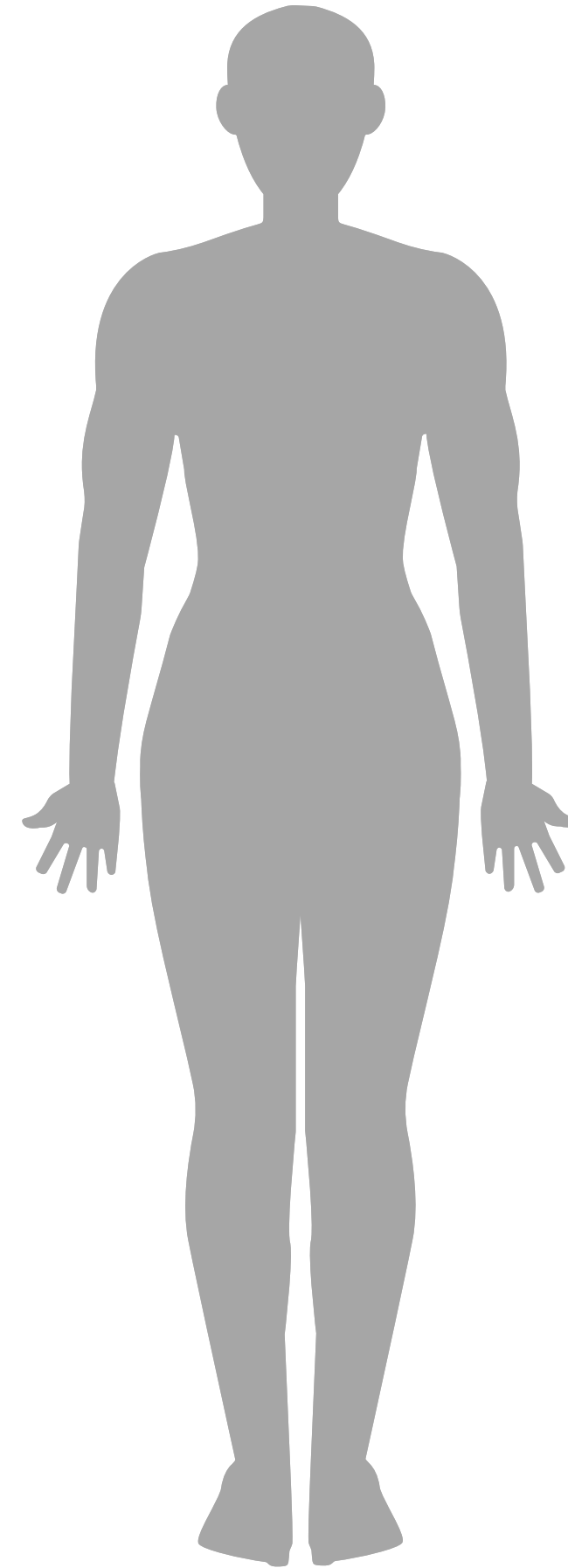
Datenbanken



Sprachschätze



Wörter für Körperteile



Sprachvergleich



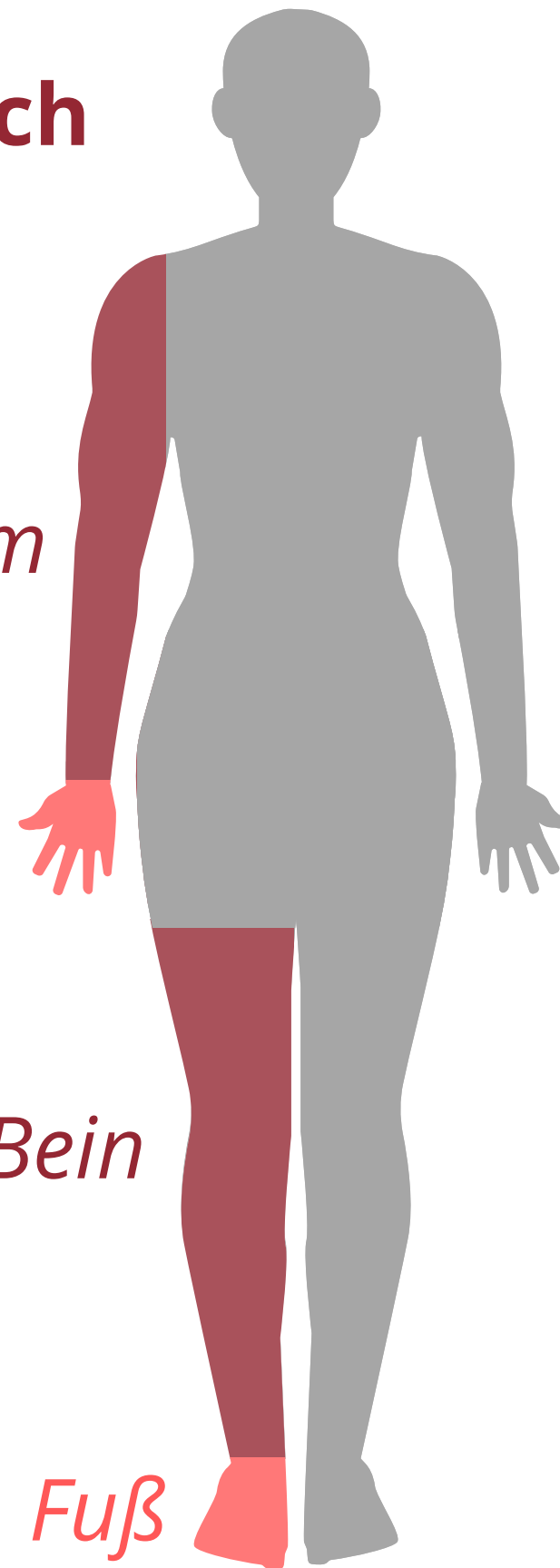
Deutsch

Arm

Hand

Bein

Fuß



Sprachvergleich



Deutsch

Tschechisch

Arm

ruka

Hand

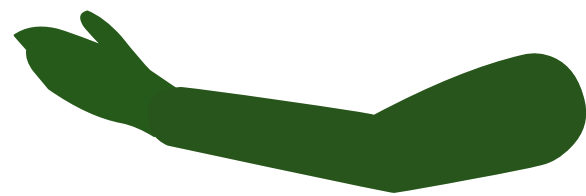
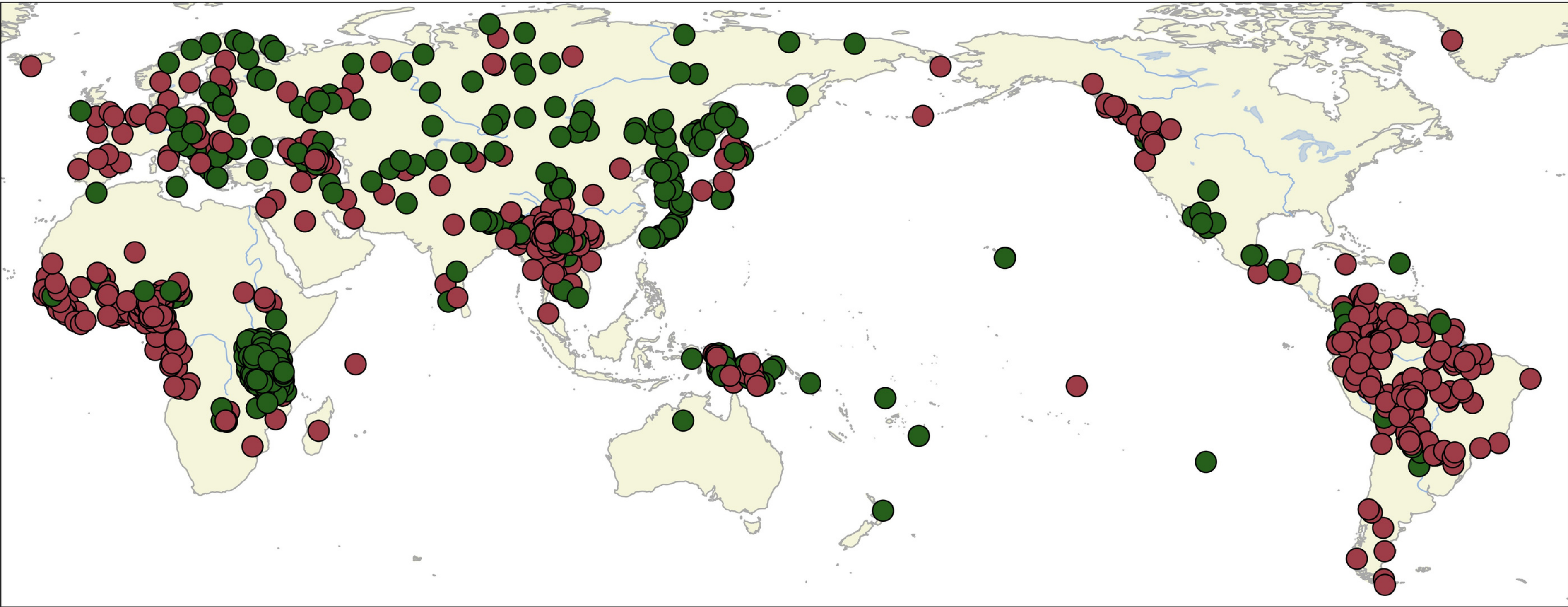


Bein

noha

Fuß





HAND/ARM ein
Wort



HAND und ARM
zwei Wörter

Warum?

Wahrnehmung

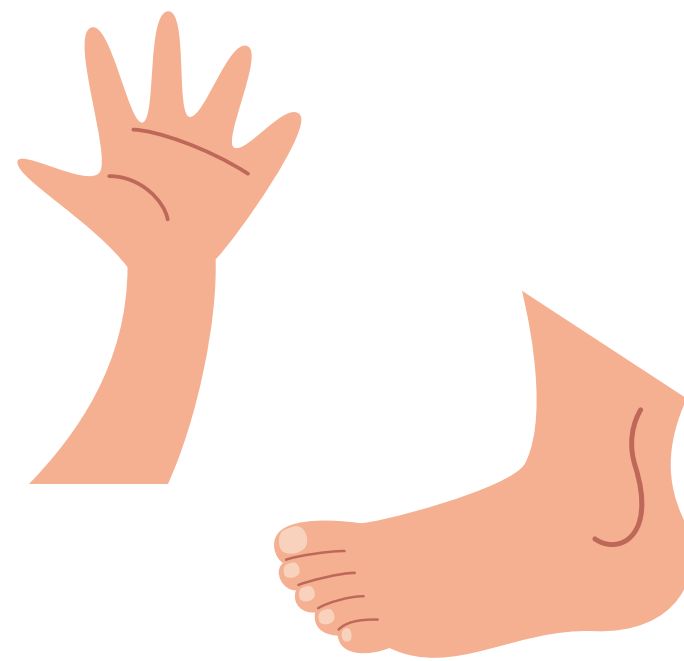


Nebeneinander

Wahrnehmung



Nebeneinander

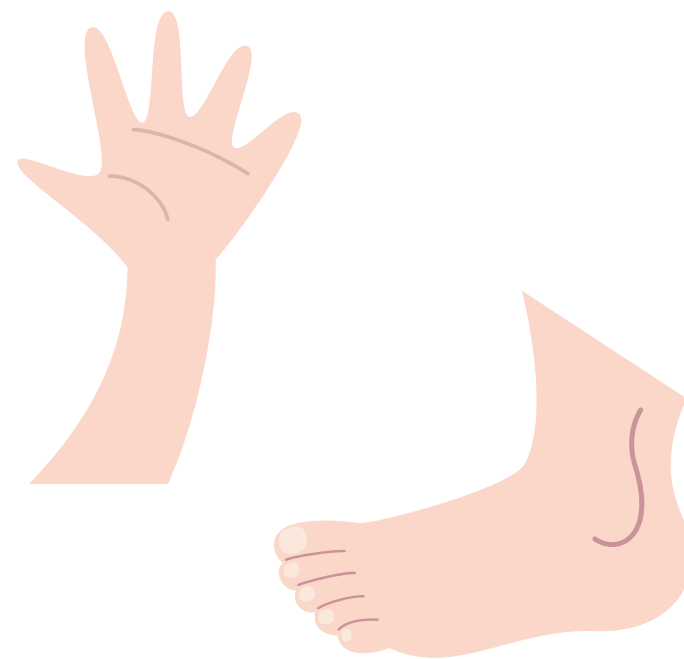


Funktion

Wahrnehmung



Nebeneinander



Funktion



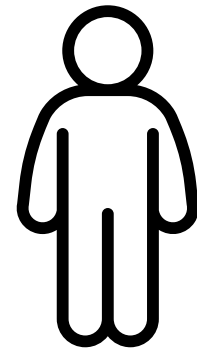
Form



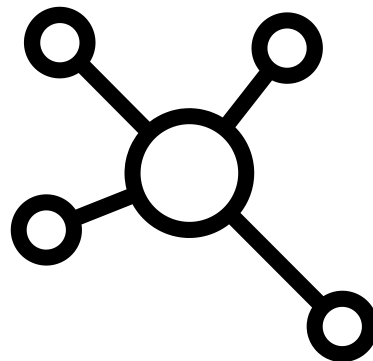
Studie



> 1.000 Sprachen

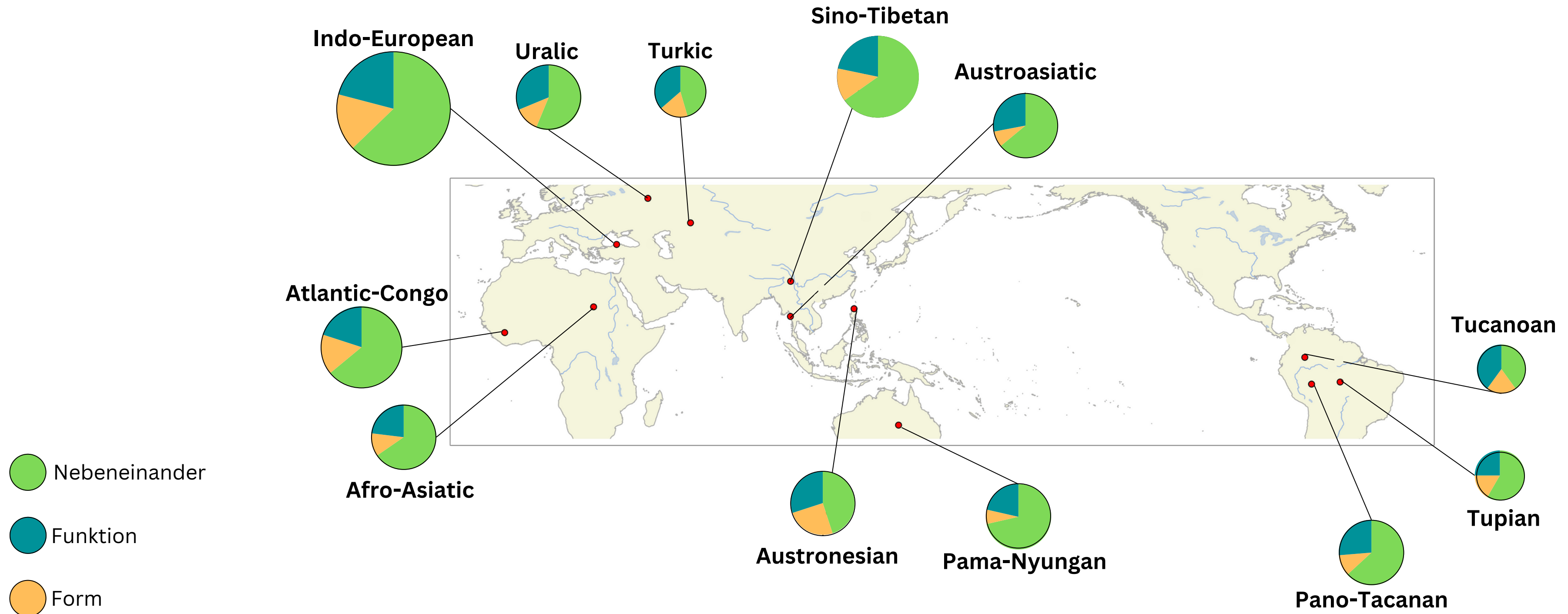


110 Körperteile

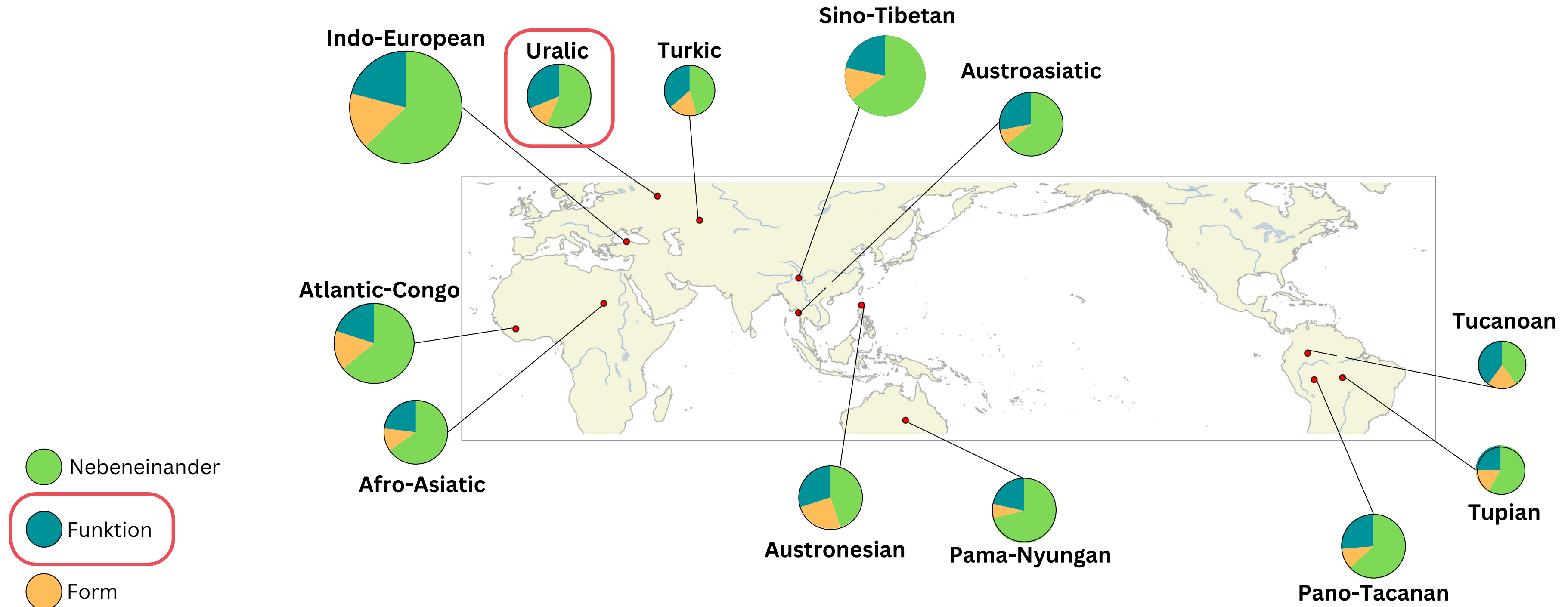


Datenanalyse

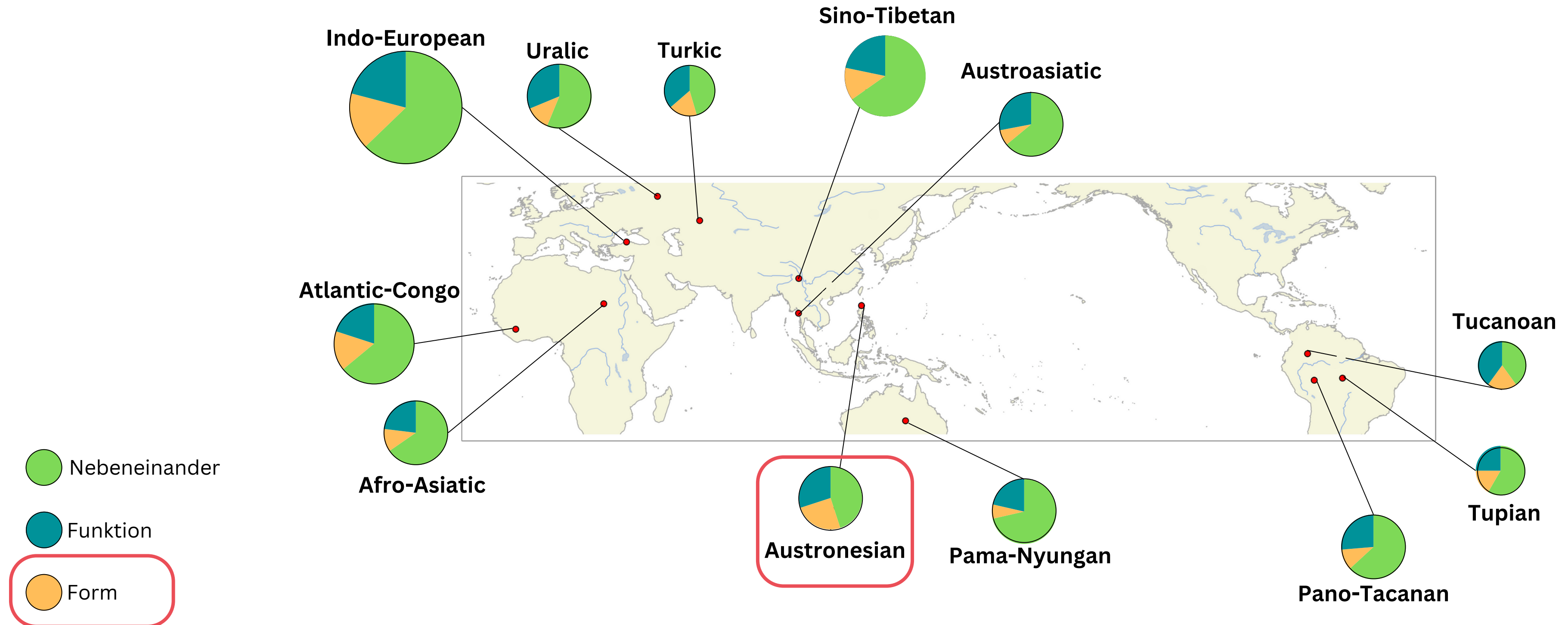
Verteilung



Verteilung



Verteilung



Zusammenfassung

Datenbanken als Schatztruhe

- Sprachvielfalt

Zusammenfassung

Datenbanken als Schatztruhe

- Sprachvielfalt
- Sprachdokumentation

Zusammenfassung

Datenbanken als Schatztruhe

- Sprachvielfalt
- Sprachdokumentation
- Sprachvergleich

Zusammenfassung

Datenbanken als Schatztruhe

- Sprachvielfalt
- Sprachdokumentation
- Sprachvergleich

Benennung von Körperteilen universell und variabel

- Nebeneinanderliegende Körperteile werden sprachübergreifend unter demselben Wort zusammengefasst.

Zusammenfassung

Datenbanken als Schatztruhe

- Sprachvielfalt
- Sprachdokumentation
- Sprachvergleich

Benennung von Körperteilen universell und variabel

- Nebeneinanderliegende Körperteile werden sprachübergreifend unter demselben Wort zusammengefasst.
- Sprachvariation entsteht durch Präferenzen für die Wahrnehmungsmerkmale Funktion oder Form.



Danke

Gracias

Thank you

Merci

Grazie

Merim

omni

omni

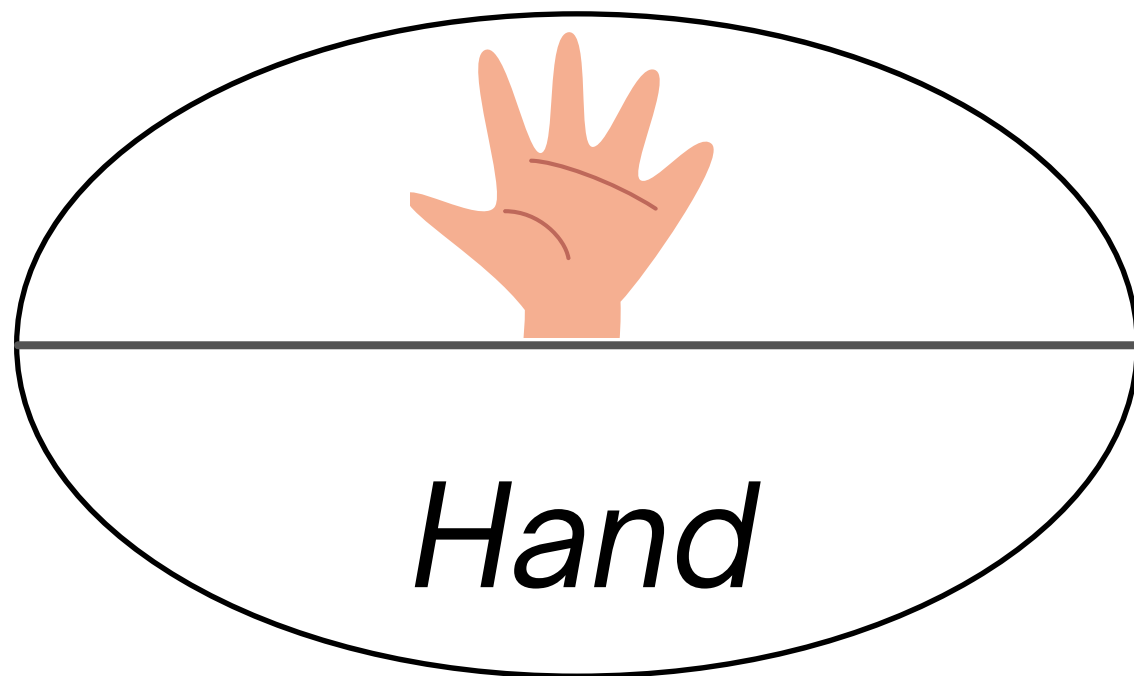
謝謝

ala

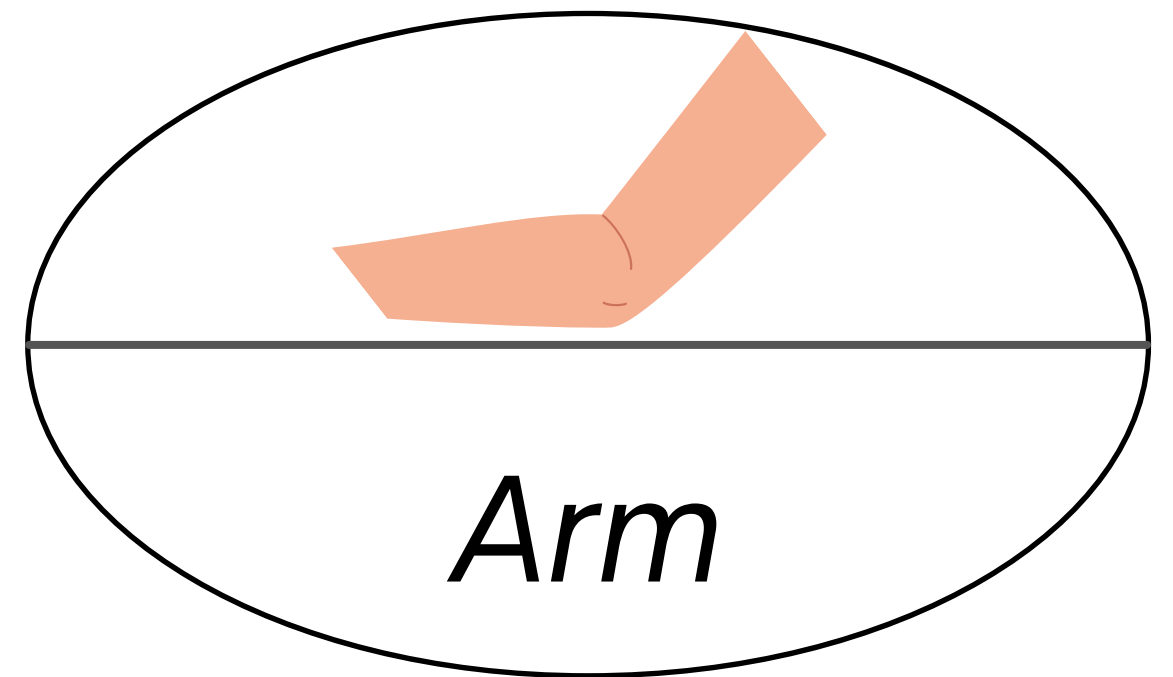
Koszona

ank

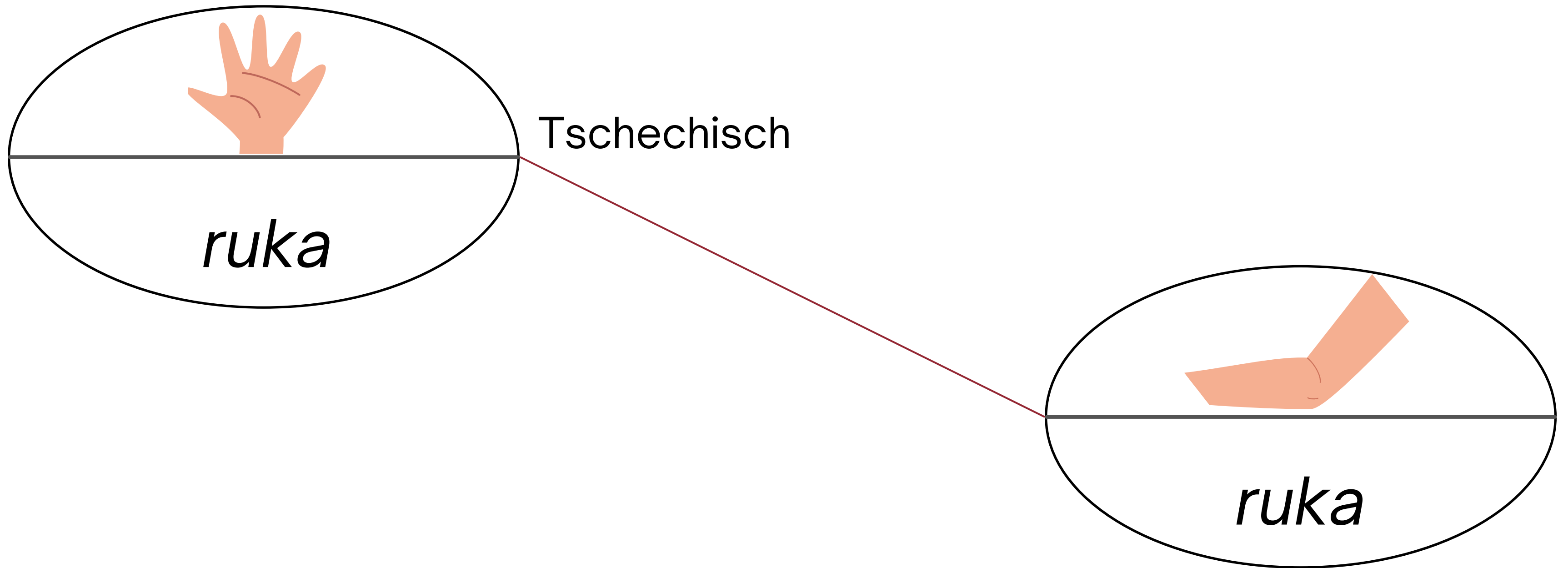
Netzwerk



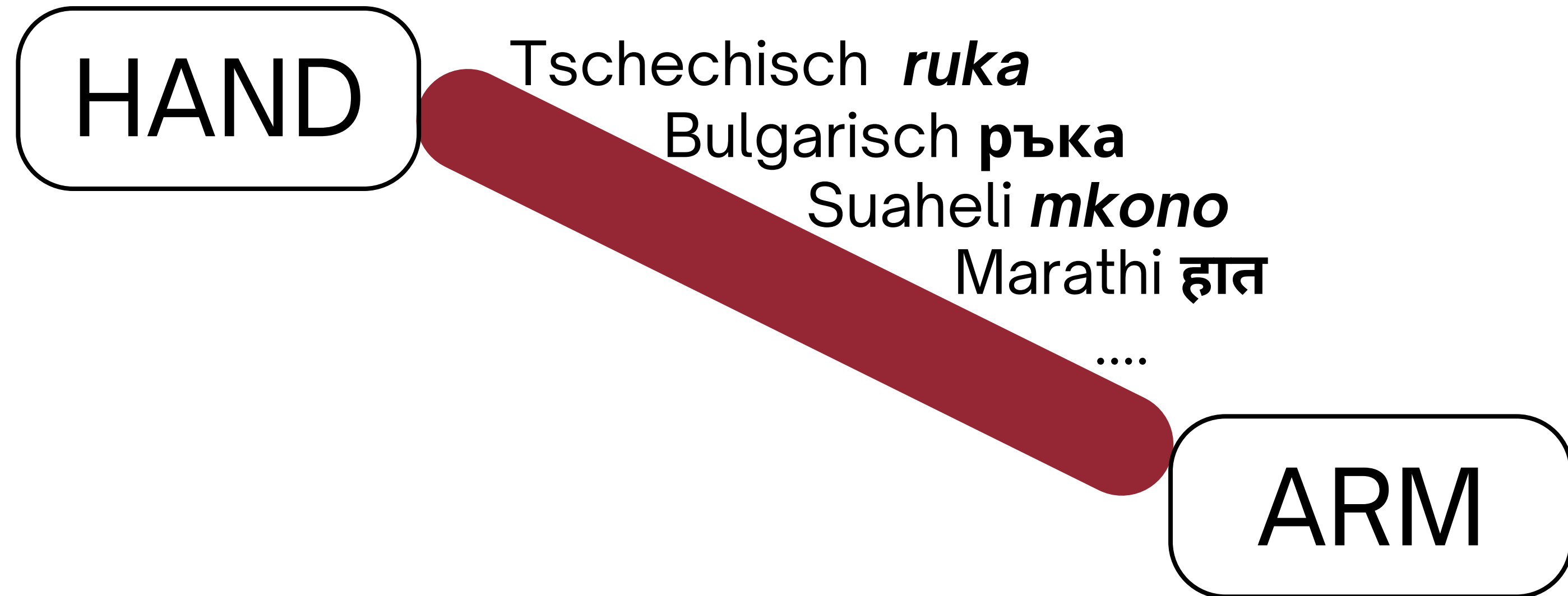
Deutsch



Netzwerk



Netzwerk



Weltweit

