

The body as a source for object names: A study of partial colexifications across languages

Annika Tjuka¹ & Johann-Mattis List²

¹ Max Planck Institute for Evolutionary Anthropology, annika_tjuka@eva.mpg.de ² University of
Passau, mattis.list@uni-passau.de

Keywords: Lexical Typology, Body Parts, Colexifications

Expressions in which a body part term refers to an object occur in most languages. Although English does not systematically use body part terms for objects (Lakoff & Johnson, 1980), numerous examples exist: *foot of the mountain*, *clock face*, etc. In addition, some languages use body part terms consistently for object descriptions (Levinson, 1994). A cross-linguistic approach that considers diverse languages is essential to identify constraints on linguistic variation (Croft, 2016). Thus, the study of cross-linguistic *colexifications* (François, 2008) offers important insights into cognitive principles that facilitate the use of body part terms for objects. As of yet, large-scale studies across many languages (e.g., Jackson et al., 2019; Xu et al., 2020; Brochhagen & Boleda, 2022) have focused on full colexifications in which the same lexical item is used for two different meanings, for example, Czech *ruka* ‘hand, arm’. Partial colexifications such as *river mouth* where one part of the word, i.e., *mouth*, is colexified, are examined primarily in studies focusing on a smaller number of languages (e.g., Schapper, 2022; Urban, 2022). In this talk, we present a new methodology that allows us to infer full and partial colexifications between body and object concepts automatically (cf. List, 2023).

The material for our study is based on a large lexical data collection, Lexibank (List et al., 2022). We selected concepts representing human body parts and everyday objects from the Concepticon reference catalog that consists of cross-linguistically comparable concepts (List et al., 2016; Tjuka et al., 2022). The 45 body concepts included, for example, HEAD, NOSE, ARM, and BONE. The 65 object concepts were from the semantic field of artifacts, landscapes, plants, food, and household items. The concepts are linked to word lists such as the Intercontinental Dictionary Series (Key & Comrie, 2021) which are curated in Lexibank (List et al., 2022). By employing the methods described in Rzymiski et al. (2020), we automatically generated a colexification network of the data. However, this preliminary network only provided full colexifications, i.e., two different concepts are linked to identical lexical items. Thus, we established a new method that detects common substrings across lexical items to extend the analysis to partial colexifications (List, 2023). This method recognizes whether a concept is expressed, for example, by a compound as in *river mouth*, where only one part colexifies with a body concept.

We analyzed 93 colexifications between a body and object concept across 997 languages from 87 language families. The results indicate that only a few body-object colexifications are frequent across several languages, but many distinct body-object colexifications exist in diverse languages. For example, the colexification between SKIN and BARK is one of the most widespread and occurs in 128 languages. There are also languages like Maori or Abui that use the partial colexification ‘tree skin’ for the concept BARK. The body-object colexifications arise from similarities based on visual and functional perception between body parts and objects. For example, shape leads to the colexification between NOSE and CAPE and spatial alignment is the basis for HEAD and MOUNTAIN SUMMIT. The study also supports previous findings that languages have preferences for certain similarities (Tilbe, 2017).

The implications of the findings are far-reaching in that they offer new insights into meaning extensions of body part terms from a cognitive perspective. In addition, the new methodological approach allows for a study of different morphological structures in the formation of cross-linguistic colexifications and can be extended to other semantic domains.

References

- Brochhagen, Thomas & Gemma Boleda. 2022. When do languages use the same word for different meanings? The Goldilocks principle in colexification. *Cognition* 226. 1–8. doi:10.1016/j.cognition.2022.105179.
- Croft, William. 2016. Typology and the future of Cognitive Linguistics. *Cognitive Linguistics* 27(4). 587–602. doi:10.1515/cog-2016-0056.
- François, Alexandre. 2008. Semantic maps and the typology of colexification: Intertwining polysemous networks across languages. In Martine Vanhove (ed.), *From polysemy to semantic change: Towards a typology of lexical semantic associations*, vol. 106 Studies in Language Companion Series, 163–215. Amsterdam/Philadelphia: John Benjamins Publishing. doi:10.1075/slcs.106.09fra.

- Jackson, Joshua Conrad, Joseph Watts, Teague R. Henry, Johann-Mattis List, Robert Forkel, Peter J. Mucha, Simon J. Greenhill, Russell D. Gray & Kristen A. Lindquist. 2019. Emotion semantics show both cultural variation and universal structure. *Science* 366. 1517–1522. doi:10.1126/science.aaw8160.
- Key, Mary Ritchie & Bernard Comrie. 2021. *The Intercontinental Dictionary Series*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Lakoff, George & Mark Johnson. 1980. *Metaphors we live by*. Chicago, Illinois: University of Chicago Press.
- Levinson, Stephen C. 1994. Vision, shape, and linguistic description: Tzeltal body-part terminology and object description. *Linguistics* 32(4-5). 791–855. doi:10.1515/ling.1994.32.4-5.791.
- List, Johann-Mattis. 2023. Inference of partial colexifications from multilingual wordlists. doi:10.48550/arXiv.2302.00739.
- List, Johann-Mattis, Michael Cysouw & Robert Forkel. 2016. Concepticon: A resource for the linking of concept lists. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk & Stelios Piperidis (eds.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation*, 2393–2400. Portorož, Slovenia: European Language Resources Association.
- List, Johann-Mattis, Robert Forkel, Simon J. Greenhill, Christoph Rzymiski, Johannes Englisch & Russell D. Gray. 2022. Lexibank, a public repository of standardized wordlists with computed phonological and lexical features. *Scientific Data* 9(1). 316. doi:10.1038/s41597-022-01432-0.
- Rzymiski, Christoph, Tiago Tresoldi, Simon J. Greenhill, Mei-Shin Wu, Nathanael E. Schweikhard, Maria Koptjevskaja-Tamm, Volker Gast, Timotheus A. Bodt, Abbie Hantgan, Gereon A. Kaiping, Sophie Chang, Yunfan Lai, Natalia Morozova, Heini Arjava, Nataliia Hübler, Ezequiel Koile, Steve Pepper, Mariann Proos, Briana Van Epps, Ingrid Blanco, Carolin Hundt, Sergei Monakhov, Kristina Pianykh, Sallona Ramesh, Russell D. Gray, Robert Forkel & Johann-Mattis List. 2020. The Database of Cross-Linguistic Colexifications, reproducible analysis of cross-linguistic polysemies. *Scientific Data* 7(1). 1–12. doi:10.1038/s41597-019-0341-x.
- Schapper, Antoinette. 2022. Baring the bones: The lexico-semantic association of bone with strength in Melanesia and the study of colexification. *Linguistic Typology* 26(2). 313–347. doi:10.1515/lingty-2021-2082.
- Tilbe, Timothy James. 2017. *Parts and wholes in Mesoamerican language and cognition*. Buffalo, New York: State University of New York at Buffalo dissertation.
- Tjuka, Annika, Robert Forkel & Johann-Mattis List. 2022. Curating and extending data for language comparison in Concepticon and NoRaRe [version 1; peer review: awaiting peer review]. doi:10.12688/openreseurope.15380.1.
- Urban, Matthias. 2022. Red, black, and white hearts: ‘heart’, ‘liver’, and ‘lungs’ in typological and areal perspective. *Linguistic Typology* 26(2). 349–374. doi:10.1515/lingty-2021-2081.
- Xu, Yang, Khang Duong, Barbara C. Malt, Serena Jiang & Mahesh Srinivasan. 2020. Conceptual relations predict colexification across languages. *Cognition* 201. 1–9. doi:10.1016/j.cognition.2020.104280.