Types of classifier systems

Marcin Kilarski[1] · Marc Allassonnière-Tang[2]

Presented at the 2024 Annual Meeting of the Linguistic Society of America (5 January - 7 January, 2024)

[1]Adam Mickiewicz University in Poznań, Poland

[2]Muséum National d'Histoire Naturelle, Université Paris Cité, Centre National de la Recherche Scientifique, France

kilarski@amu.edu.pl, marc.allassonniere-tang@mnhn.fr

Motivation

Diverse types of classifiers are a prototypical example of linguistic diversity and the capacity of the human mind for categorization.

Gap in research

Descriptions of classifiers typically rely on small-scale surveys or case studies of classifier languages. In WALS [1] and WACL [2], we only find numeral classifiers.

Proposed solution

Constructing a database of classifier types in the world's languages and determining the distribution of semantic values in classifier languages.

Examples of classifier types

Numeral classifier (Mandarin)
yi4 zhi1 gou3
one CLF.ANIM dog
'one dog' (anim = animal)

Possessive classifier (White Hmong)
nws rab riamntaj
he CLF.INST sword
'his sword' [3] (inst = instrument)

Noun classifier (Zhuang)
tu2 mou1 kwn1 bou3 im5
CLF.ANIM pig eat not enough
'The pig is not full.' [4]

Deictic classifier (Kadiwéu)
i-n:i-wa-tale gonele:gi-wa-di
MASC-CLF.NXT-PL-two man-N-PL
'two men' [5] (nxt = non-extended)

Materials

The DReaM corpus [6]: Grammars and grammar sketches written in English, currently 7126 source documents describing 3240 languages.

Research questions

Underlying principles of categorization in classifier systems

- What are classifiers and what types of classifier systems are distinguished?
- What semantic values are found in classifier systems and how are they structured?

Universal vs. language-specific distribution of semantic values and the interaction between semantics and types of classifier systems

- What is the distribution of semantic values and types of classifiers?
- Is there a preference among classifier types for certain semantic values?

Potential pitfalls

The diversity of terms, e.g., 'projectives', 'company words', 'quantifiers'. The fuzzy distribution of noun classifiers. Manual checking will be conducted.

Preliminary output

For all the sources for each language, we automatically extracted the preceding word for each occurrence of the term 'classifier', e.g., 'numeral classifier'. The preliminary results show that corpora combined with NLP methods have a high potential for identifying classifier types in the world's languages.

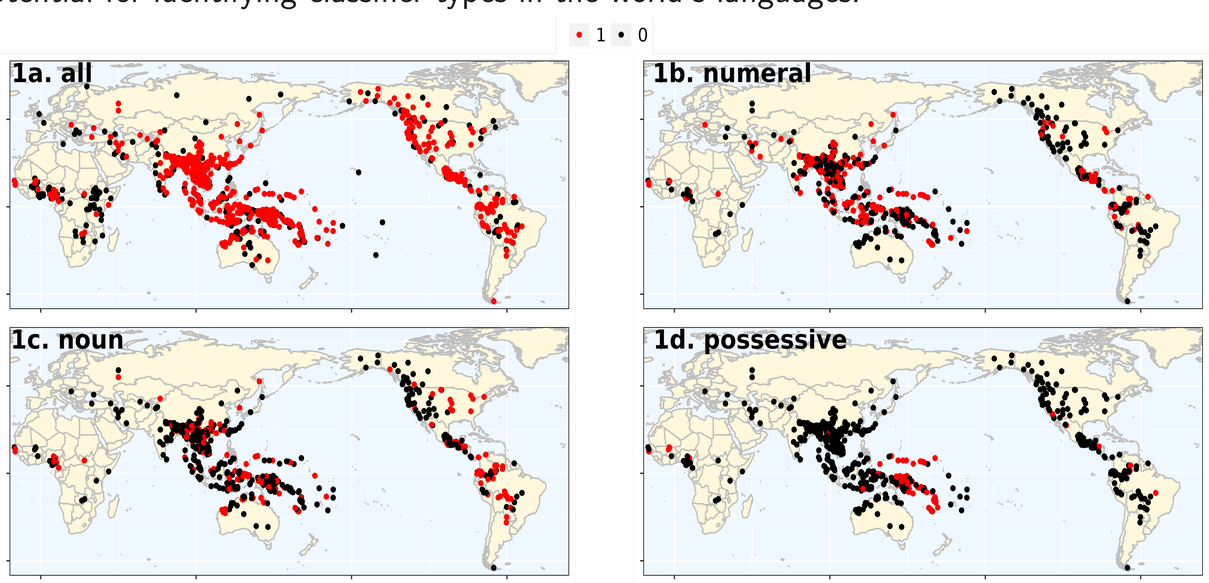

Figure 1. In the preliminary study, among 986 languages, there are 651 (66.02%) classifier languages (red = classifier languages, black = languages without classifiers).

Acknowledgements

The project is funded by the National Science Centre, Poland (UMO-2022/47/B/HS2/02999).

References

[1] Gil, D. 2013. Numeral classifiers. In M.S. Dryer & M. Haspelmath (eds.), The world atlas of language structures online. Leipzig: Max Planck Institute for Evolutionary Anthropology. https://wals.info/chapter/55.

[2] Her, O.-S., H. Hammarström & M. Allassonnière-Tang. 2022. Defining numeral classifiers and identifying classifier languages of the world. Linguistics Vanguard 8(1). 151-164. https://doi.org/10.1515/lingvan-2022-0006.

[3] Bisang, W. 1988. Eine Auswahl mit Interlinearübersetzung aus Jean Mottin, Contes et Légendes Hmong Blanc. Arbeiten des Seminares für allgemeine Sprachwissenschaft der Universität Zürich Nr. 8.

[4] Qin, X. 2007. Concurrent functions of Hawyiengz Zhuang classifiers. Mon-Khmer Studies 37. 165-178.

[5] Sandalo, F & Michelioudakis, D. 2016. Classifiers and Plurality: Evidence from a deictic classifier language. The Baltic International Yearbook of Cognition, Logic and Communication 11. 1-40. https://doi.org/10.4148/1944-3676.1112

[6] Virk, S.M., H. Hammarström, M. Forsberg & S. Wichmann. 2020. The DReaM Corpus: A multilingual annotated corpus of grammars for the world's languages. Proceedings of the Twelfth Language Resources and Evaluation Conference, 878-884. Marseille: European Language Resources Association. https://aclanthology.org/2020.lrec-1.110.