

# Animacy effects in Yucatec Maya: a preliminary survey of animacy-related data in the corpus CoCoYum



Eric Thurau

Mayan Languages

Elisabeth Verhoeven, Rodrigo Gutiérrez-Bravo, Stavros Skopeteas

Institut für deutsche Sprache und Linguistik

## Research question:

How are the findings of Bohnemeyer & Butler (2024) reflected in corpus data? Does the data show similar strategies being employed with regards to animacy effects? And what can the data tell us about possible processing effects in Yucatec Maya?

### Background & Motivation:

- Previous engagement with non-structural animacy effect research; most of which was based on non-head-marking languages
- Starting point: the findings of the elicitation study of Bohnemeyer, Butler (2024); using a corpus to see how these insights are borne out in a large collection of Yucatec Maya data
- Special focus on resolution of same rank clashes (animate vs. animate), expectation: dominance of passive constructions
- Consideration of data within processing frameworks: the surprisal approach (Macdonald et al., 2020) and disambiguation (Chen, Hale, 2021)

### Methodology:

- Initial review of the CoCoYum-Corpus (Lehmann, 2021) using the corpus-based search engine ANNIS
- Comparison of passive vs. left dislocation (via topicalization) occurrences
- Drafting animacy related categories for nouns (animals, vehicles) and verbs (transitive, semantic relation with animacy)
- Systematic search of the established categories and classification by representation type (norm, passive, dislocation of A/P, intransitive, other)
- Frequency analysis & graphical representation using R Studio (& packages `cairo`, `dplyr`, `ggplot2`)

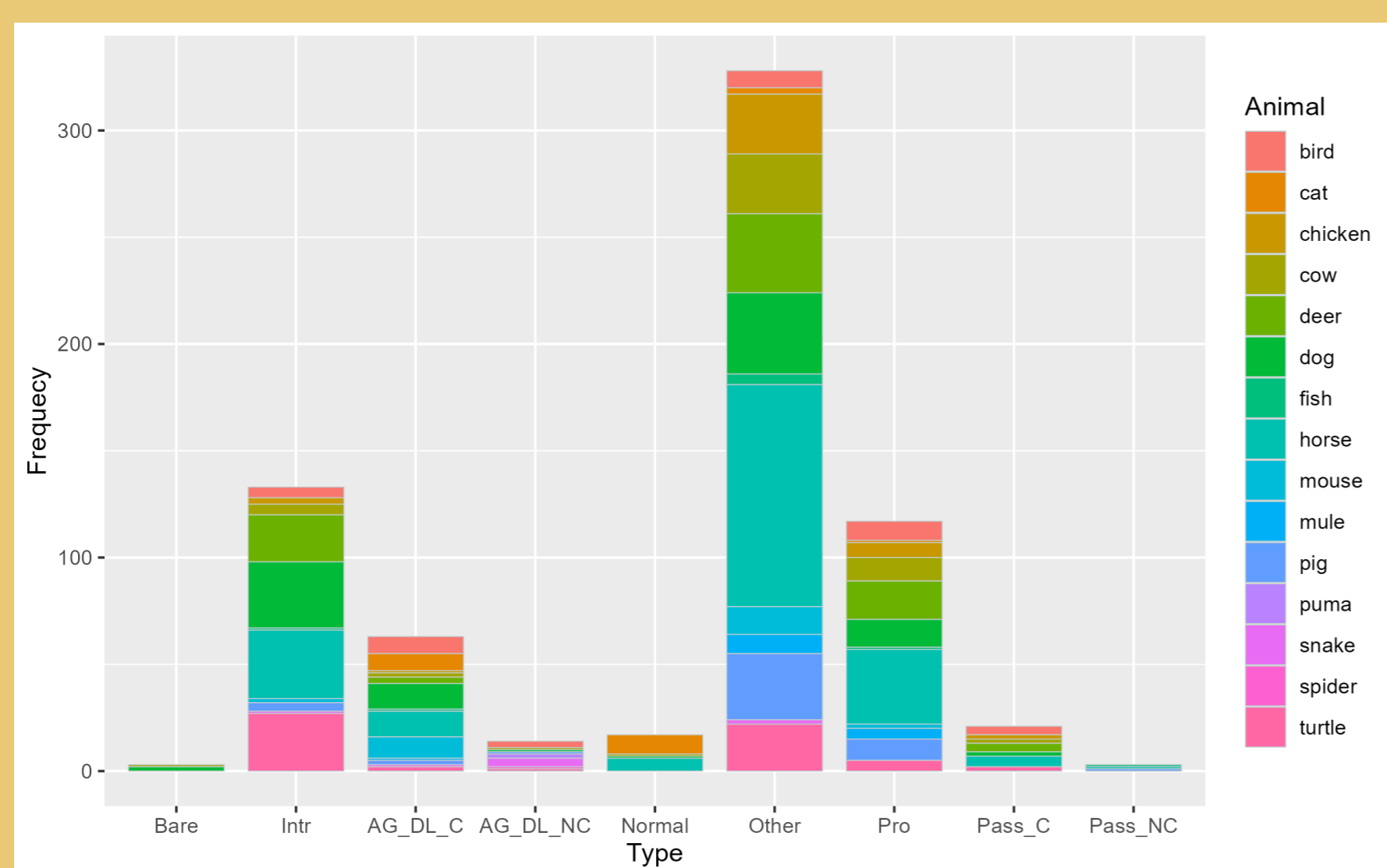


Figure 1: frequency of sentence types according to animal mentions

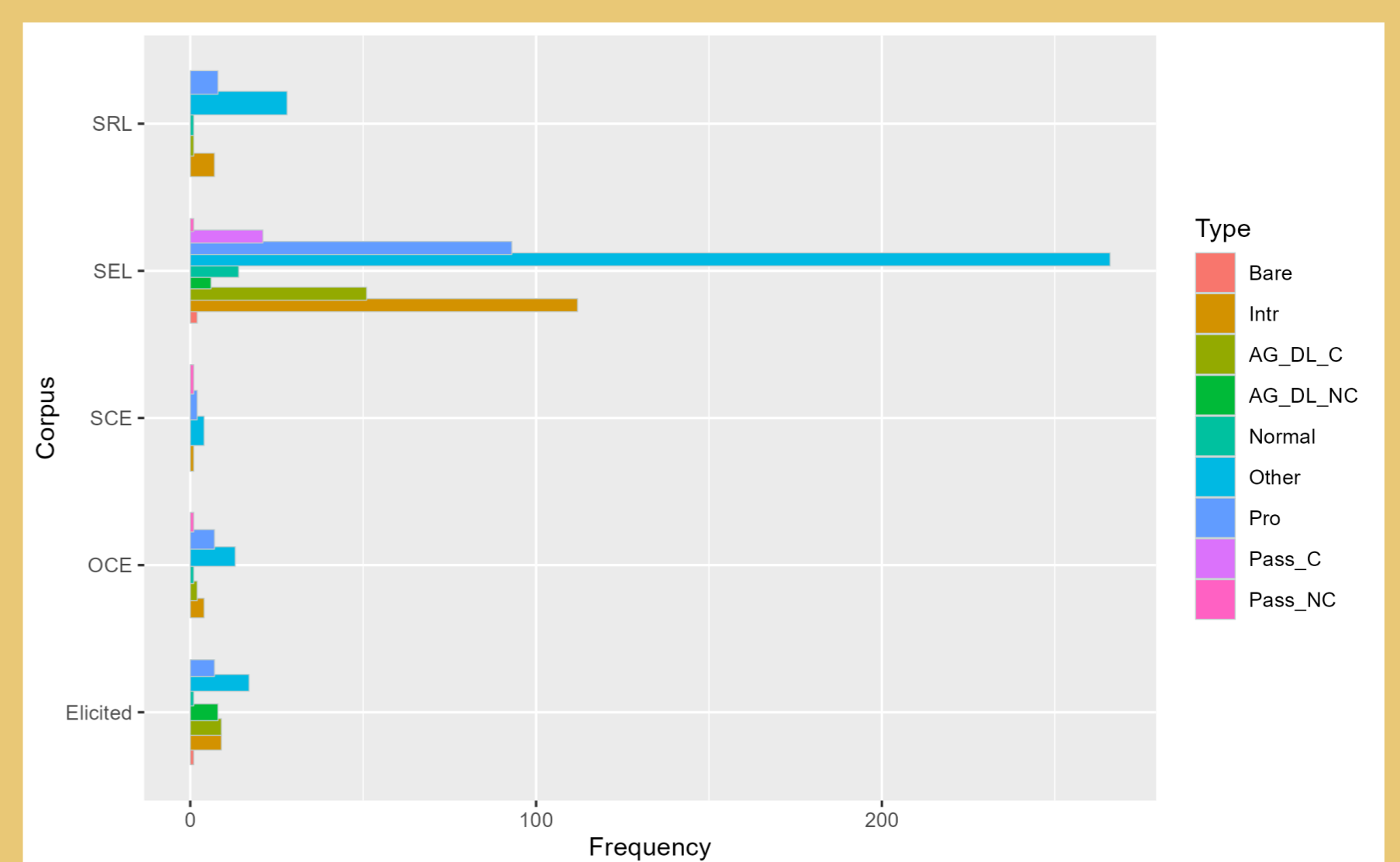


Figure 2: frequency of sentence types by subcorpus

### Results:

- Uneven distribution of the data: out of 169.000 tokens only about 700 refer to animals (non-pronominal form), vehicle-mention < 10; similarly sparse results for the established verb list (202, of which ca. 58% derive from the verb *kill*)
- Most data comes from the SEL subcorpus, containing edited version of narratives, procedural texts, etc.
- Largely, animal mentions belong to one of three sentence types: intransitives, bare cross-references, other (complex clauses, imperatives, adverbials)
- For animacy clashes (animate vs. animate), left dislocations of the agent were more common than the use of passive constructions
- In general, passive constructions rarely appeared in the established categories (23 occurrences out of about ca. 1000 in the entire corpus)
- One occurrence of the patient left dislocation (for the verb *kill*):

(1) *iíhoh ts'o'k u kúins -ik -echhun -túul = o'*  
 son TERM SBJ.3 kill -INCMPL -ABS.2.SG one -CL.AN = D2  
 Son, one [of my daughters] has killed you

- Possible prevalence of agent left dislocation as a result of disambiguation and/or to reduce mental load, in line with the bare cross-referential type

### Problems and future possibilities:

- Sparse dataset allows for almost no inferences
- Difficulty in establishing salient categories w.r.t. the question at hand; catch-all category „other“ hard to handle with insufficient language skills
- Despite concerns with data: results don't seem to confirm with initial expectations
- In the future: possibility to use similar kinds of corpus research to model information-theoretic approaches, i.e. grammatical models based on head-marking to test for disambiguation

### Literature:

Aissen, J. (1997). On the syntax of obviation. *Language* 73 (4): 705–750.  
 Bohnemeyer, Jürgen 2024. Animacy and topicality in Yucatec sentence production. Zenodo. <https://doi.org/10.5281/zenodo.10578820>. In Gutiérrez Bravo, Rodrigo, Stavros Skopeteas, Elisabeth Verhoeven (eds.), *Glottothèque: Mayan Languages*. Berlin/Göttingen/Mexico City: online resource.  
 Bohnemeyer, J. & T. J. Tilbe. (2021). Argument realization and discourse status in Yucatec, a purely head-marking language. *Amerindia* 43: 249–289.  
 Chen, Z. and Hale, J.T. (2021). Quantifying Structural and Non-structural Expectations in Relative Clause Processing. *Cogn Sci*, 45: e12927. <https://doi.org/10.1111/cogs.12927>  
 Lehmann, Christian. 2021. *Yuc\_text.db: A collection of Yucatec texts and examples*. In Elisabeth Verhoeven, Nico Lehmann, Frederic Blum (eds.), *Collective corpus of Yucatec Maya (CoCoYum)* [Insert Version number]. Berlin: Humboldt-Universität zu Berlin. (Available online at <https://scm.cms.hu-berlin.de/lehmannx/cocoyum-public>, accessed on [28.03.2024])  
 Macdonald, R., Brandt, S., Theakston, A., Lieven, E. and Serratrice, L. (2020). The Role of Animacy in Children's Interpretation of Relative Clauses in English: Evidence From Sentence–Picture Matching and Eye Movements. *Cogn Sci*, 44: e12874. <https://doi.org/10.1111/cogs.12874>