Cross-linguistic patterns in Person systems reflect efficient coding

Mora Maldonado,* [1] Noga Zaslavsky,* [2] and Jennifer Culbertson[1]
[1]Centre for Language Evolution, University of Edinburgh; [2]Department of Brain and Cognitive Sciences and Center for Brains Minds and Machines, MIT; * Equal contribution.

Person systems refer to individuals as a function of their conversational role: there is a speaker (e.g., 'I'), an addressee (e.g., 'you'), and others (e.g., 'they'). Like other semantic domains, person systems exhibit constrained cross-language variation (Cysouw, 2003). For example, while many languages express the *you and us* inclusive meaning as a form of first person (1st-inclusive, e.g., 'we'), Zwicky (1977) observed that no known language expresses that meaning as a form of second person (2nd-inclusive), which suggests an asymmetry in the representation of the speaker and addressee. Current linguistic theories account for this by positing strong grammatical constraints on possible systems (Harbour, 2016). However, a recent study (Maldonado and Culbertson, 2020) challenged this view by showing that the unattested 2nd-inclusive system is learnable in artificial settings, while the unattested 3rd-inclusive system is not. This finding is not explained by the aforementioned theories, leaving open the question of why these cross-linguistic patterns emerge.

Here, we address this open question by testing an alternative, information-theoretic hypothesis (Zaslavsky et al., 2018), which argues that languages efficiently encode meanings into words by optimizing the Information Bottleneck (IB: Tishby et al., 1999) tradeoff between the complexity and accuracy of the lexicon. This approach is grounded in Rate–Distortion theory (RDT: Shannon, 1948), and has gained empirical support in several semantic domains, e.g, color and containers. It is also closely related to other notions of efficiency (Kemp et al., 2018) that are not grounded in RDT but have been applied to domains such as kinship and indefinites (Kemp et al., 2018; Denic et al., 2020), which are qualitatively more similar to person. Therefore, the person domain poses an important test case for the applicability of RDT to the lexicon.

First, we show that the framework of Zaslavsky et al. (2018) allows us to formulate an 'egocentric' bias towards a distinct representation of the speaker, and test the proposal that Zwicky's observation stems from this bias (Maldonado and Culbertson, 2020). Specifically, we derive two compression models: (i) an egocentric model that predicts that languages efficiently encode the domain in the presence of this bias; and (ii) an unbiased model that predicts that languages efficiently encode the domain given that all entities are equally salient. If an egocentric bias shapes person systems, in addition to pressure for efficiency, then the egocentric model should provide a better account of our data and distinguish between attested and unattested systems. For each model, we computed the IB theoretical limit of efficiency, defined by the set of optimal systems for different complexity–accuracy tradeoffs. We also evaluated the tradeoffs attained by ten commonly attested person systems, the two unattested systems mentioned above, and 1,500 hypothetical systems. The results show that the attested systems are near-optimally efficient according to the egocentric model, in contrast to most hypothetical systems. In addition, the egocentric model predicts a substantial efficiency gap between the attested and unattested systems, whereas the unbiased model predicts that the unattested systems are as efficient as attested systems. This suggests that Zwicky's observation may be explained by functional pressure for efficient coding in the presence of an egocentric bias, and this explanation is consistent with the findings of Maldonado and Culbertson (2020). Finally, an initial analysis of a larger typological dataset (Cysouw, 2003) suggests that our result generalize well across languages.

This work shows that person systems across languages achieve near-optimal compression, providing converging evidence for the applicability of RDT to the lexicon. Furthermore, it suggests a principled way to study how cognitive biases may influence the lexicon, and may explain typological tendencies, such as Zwicky's observation, which previous theories have struggled to explain.

References

Cysouw, M. (2003). *The Paradigmatic Structure of Person Marking*. OUP Oxford, Oxford, UK.

Denic, M., Steinert-Threlkeld, S., and Szymanik, J. (2020). Complexity/informativeness trade-off in the domain of indefinite pronouns. In *Proceedings of the 30th Semantics and Linguistic Theory Conference*.

Harbour, D. (2016). *Impossible Persons*. Linguistic Inquiry Monographs. MIT Press, Cambridge, MA.

Kemp, C., Xu, Y., and Regier, T. (2018). Semantic Typology and Efficient Communication. *Annual Review of Linguistics*, 4(1):109–128.

Maldonado, M. and Culbertson, J. (2020). Person of interest: Experimental investigations into the learnability of person systems. *Linguistic Inquiry*, pages 1–71.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27.

Tishby, N., Pereira, F. C., and Bialek, W. (1999). The Information Bottleneck method. In *Proceedings of the 37th Annual Allerton Conference on Communication, Control and Computing*.

Zaslavsky, N., Kemp, C., Regier, T., and Tishby, N. (2018). Efficient compression in color naming and its evolution. *Proceedings of the National Academy of Sciences*, 115(31):7937–7942.

Zwicky, A. M. (1977). Hierarchies of person. In *Proceedings from the Chicago Linguistic Society*, volume 13, pages 714–733, Chicago, MI. Chicago Linguistic Society.