Trading off robust information transmission against effort in language learning and language structure

Maryia Fedzechkina[1], Elissa L. Newport[2], T. Florian Jaeger[3]

[1]Psychology, University of Pennsylvania, [2]Neurology, Georgetown University, [3]Brain & Cognitive Sciences, University of Rochester

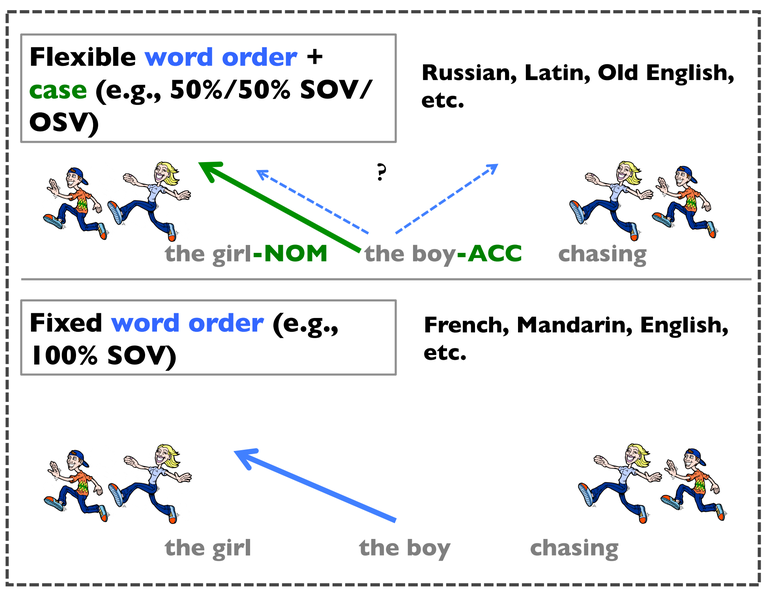
{*mfedze@sas.upenn.edu; eln10@georgetown.edu; fjaeger@bcs.rochester.edu*}

Abstract

We examine the hypothesis that at least some cross-lexical and grammatical properties of languages are beneficial for efficient information transmission [1]. We argue that these properties at least in part originate during learning. In a miniature artificial language experiment, we investigate whether learners are biased to induce grammatical systems that balance effort against uncertainty about the intended sentence meaning.

Case & word order trade-off

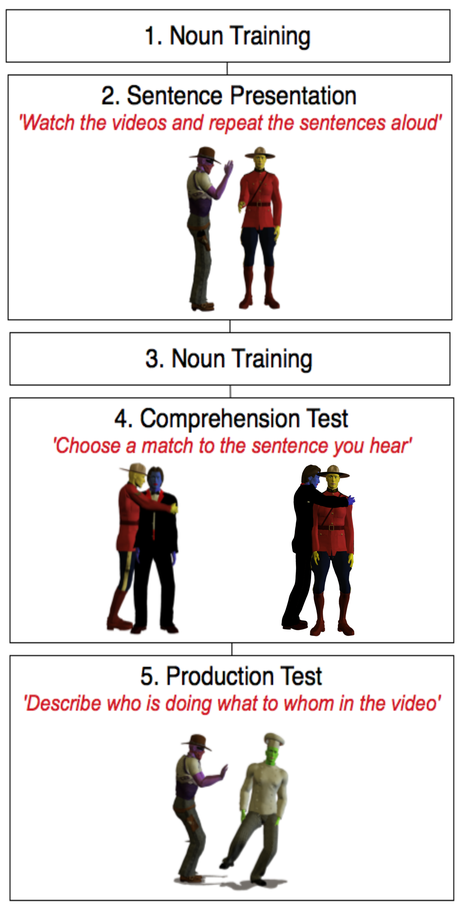
- Languages with flexible word order are more likely to have a case system cross-linguistically.
- Languages with fixed word order are less likely to have a case system cross-linguistically [2, 3].



Experiment design

A miniature artificial language learning study

- 60 monolingual English speakers
- over 3 consecutive days



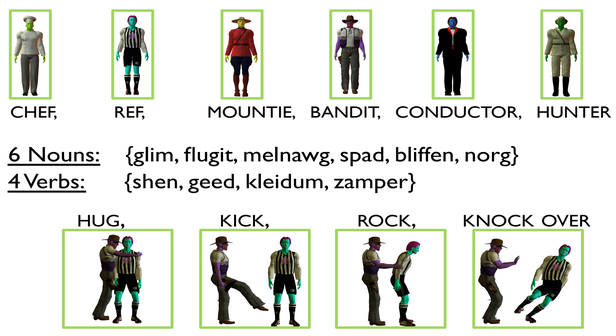
Miniature input grammars

Fixed word order language:
SOV/OSV word order 100/0%
Object case-marking 67%
(no subject case-marking)

Flexible word order language:
SOV/OSV word order 75/25%
Object case-marking 67%
(no subject case-marking)

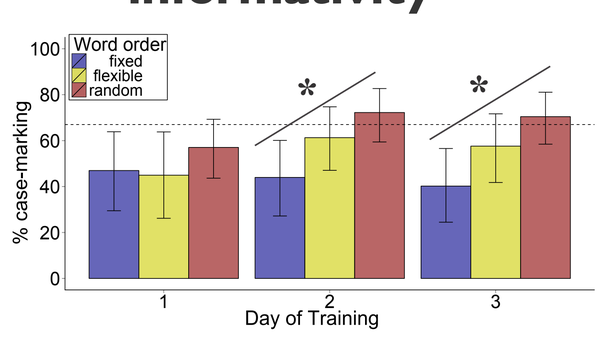
Random word order language:
SOV/OSV word order 50/50%
Object case-marking 67%
(no subject case-marking)

Miniature input lexicon

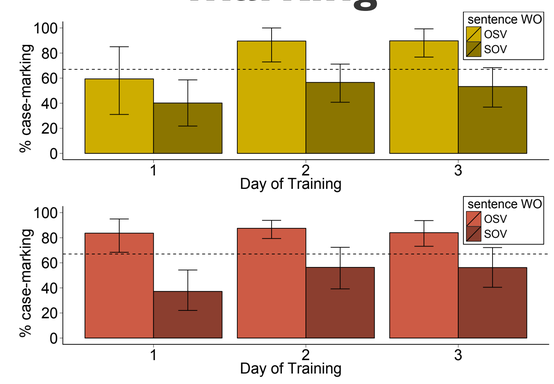


Results

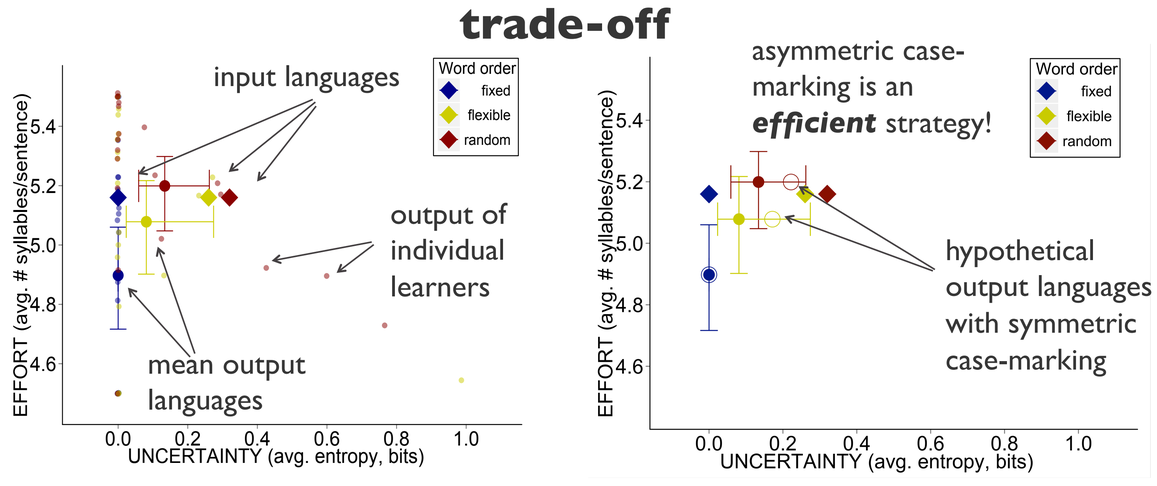
Learners' use of case increases with its informativity



Learners introduce asymmetrical case-marking



Effort & robust information transmission trade-off



- Learners were more likely to acquire and produce case the more informative it was in the input.
- Learners of the fixed order language dropped case in their productions; learners of the non-fixed order languages introduced asymmetric case-marking.
- Learning outcomes are virtually all guided by the underlying preference to balance effort and robust information transmission (59 out of 60 learners follow this principle).
- Learners differ in how they implement this principle.

Discussion

1. **Learners trade off cues to sentence meaning** during language acquisition.
2. **This behavior is not reducible to native language biases** (all participants are monolingual native speakers of English, a language with no nominal case system).
3. **Learning outcomes provide an insight into linguistic diversity.**
   - For any grammatical system, the general principle of trading off robust information transmission against effort can manifest in a variety of innovations created by learners.
4. **Learning outcomes mirror typological data**:
   - Learners drop case-marking for the fixed word order language but maintain a case system for the non-fixed order languages (e.g., Latin → French, Old Church Slavonic → Russian, Old English → Modern English).

References

[1] Piantadosi, S., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proc Natl Acad Sci USA*, 108(9), 3526

[2] Blake, B. J. (2001). *Case.* Cambridge: CUP.

[3] Sapir, E. (1921). *Language: An introduction to the study of speech.* New York: Harcourt, Brace.

Acknowledgements

This work was supported in part by NSF grants BCS-0845059 and IIS-1150028, an Alfred P. Sloan Fellowship as well as an Alfred P. Sloan fellowship to TFJ and by NIH grants DC00167 and HD037082 to ELN. We thank T. Stanley and S. Krishnan for their help with data collection and coding, C. Anzalone and G. Smith for their help with data coding; C. Donatelli, I. Minkina and A. Wood for their help with video stimuli. We thank S. Fraundorf, S. Bibyk, S. Corbin, J. Schepens, A.P. Salverda, and B. Pajak for their comments on earlier versions of this work.