PLAIN SIMPLE COMPLEX STRUCTURES: THE EMERGENCE OF OVERSPECIFICATION IN AN ITERATED LEARNING SETUP

STEFAN HARTMANN[1], PEETER TINITS[2], JONAS NÖLLE[3], THOMAS HARTMANN[4], MICHAEL PLEYER[5]

[1] German Department, University of Mainz, Jakob-Welder-Weg 18, 55128 Mainz, hartmast@uni-mainz.de; [2]Dept. of Linguistics, Tallinn University; [3]Center for Semiotics, Aarhus University; [4]Dept. of Electrical Engineering, Karlsruhe Institute of Technology; [5]English Department, Universität Heidelberg

Natural languages differ in their degree of overspecification, the extent to which overt semantic markers are required even when irrelevant in the given context. More overspecified languages have been described as typologically more complex based on user effort (Kusters, 2003) and information theory (Dahl, 2004). But how and why does systematic and obligatory overspecification emerge in the first place? While such developments could in principle be seen as mere accumulation of arbitrariness, recent research has emphasized the importance of context in the emergence of different types of language systems (e.g. Winters et al., 2015). To test if subtle contextual pressures can enforce a bias towards overspecification which eventually influences language structure, we designed an Iterated Learning experiment (cf. Kirby et al., 2014) that allowed for variation in the degree of overspecification in the individual participant's output.

205 volunteers were recruited online and allocated into chains of 5 individuals. They were then trained on an artificial language and subsequently asked to point out objects to an alien. The output of generation *n* was used as input for generation *n*+1. The initial language for all chains consisted of 4 root words (e.g. *meeb* 'ball') and 2 color markers (*pu* 'blue', *li* 'yellow') used systematically and maximally efficiently: The color marker was used only if it was required to disambiguate the object in context. Across conditions, 16 of the 32 trials required disambiguation between two objects of the same type (e.g. a blue cup and a yellow cup, see Fig. 1, left panel, bottom row). In the Control condition, only one single object was displayed in the remaining 16 trials. In the Distractor condition, by contrast, the other 16 trials consisted of pictures showing two different

objects (e.g. cup and pen). Both color (yellow vs. blue) and position (left vs. right) were assigned randomly to the objects in these pictures, such that the color of target and distractor would match in half of the trials (as in Fig. 1) and differ in the remaining trials. As the two types of referential context are very similar in the Distractor condition, while they are quite distinct in the Control condition, we predicted a higher degree of overspecification for the former condition: In these trials, overspecification relieves the language user from the need to distinguish contexts in which the same type of object is displayed twice from contexts featuring two different types of object.

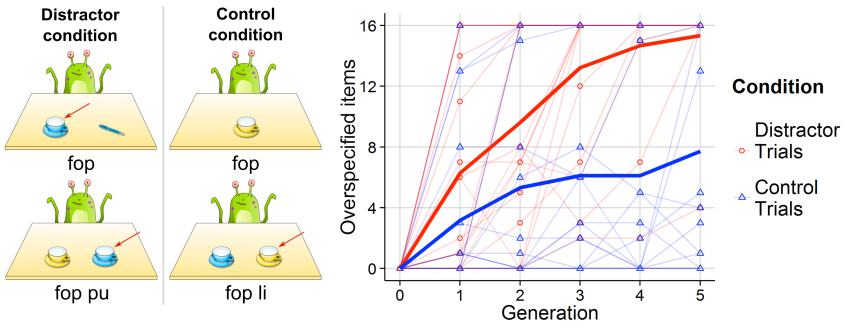

Fig. 1. Left: Experimental design. Right: Number of overspecified color markers (max. 16) for each chain across conditions. 5 chains were omitted due to noncompliance with the instructions (see supplementary materials).

Overspecification increased in both types of trials but proved more pervasive in the Distractor condition. Here, the color marker became fully obligatory in Generation 5 in 17 out of 18 chains, while it was used significantly less in Generation 5 of the control trials (two-sample $t(34)=-4.06$, $p_{\text{two-tailed}}<.001$, $r=.57$, Fig. 1).

These results indicate that contextual pressures can promote the evolution of obligatory semantic markers from adjective-like modifiers in an Iterated Learning setup. Wherever the use of color markers becomes mandatory across contexts, "color" becomes an obligatorily marked semantic distinction. Although the resulting language can be described as more complex on the typological criteria mentioned above, the use of overspecified semantic markers here is a result of strategies facilitating the task at hand in a specific context. This suggests that, given certain contextual pressures, complex structures can sometimes be the simpler solution.

Acknowledgments

The authors wish to thank James Winters for helpful comments on a previous version of the experiment reported on here. In addition, we are grateful to three anonymous reviewers for valuable comments and suggestions on the present paper.

References

Kirby, S., Griffiths, T., & Smith, K. (2014). Iterated Learning and the Evolution of Language. *Current Opinion in Neurobiology, 28*, 108–114.

Kusters, W. (2003). *Linguistic Complexity*. Utrecht: LOT.

Dahl, Ö. (2004). *Growth and Maintenance of Linguistic Complexity*. Amsterdam: John Benjamins.

Winters, J., Kirby, S., & Smith, K. (2015). Languages adapt to their contextual niche. *Language and Cognition, 7(3)*, 415–449.