An Agent-Based Model to Explain the Emergence of Dominant Word Orders

It is known that many languages today have a dominant word order. The overall distribution of word orders in the world languages is as (SOV, SVO) > VSO > (VOS, OVS) > OSV. But the fact that these word orders are not evenly distributed has been the subject of many research and where these dominant word orders came from is still a matter of curiosity. Various explanations have been given on this subject with the help of newly emerged sign languages, artificial language experiments and computational models: A dominant word order may be absent in the very initial stages of a young system; there may be cognitive biases (independent of surrounding/native languages); communicative pressures shaping word order preferences; and considering the prevalence of SOV and SVO, there is so far robust evidence for the S before O pattern, displaying a bias for the agent preceding the patient.

In this study, we propose a model to test the existing explanations regarding the emergence of dominant word orders. The model comprises three components: agent characteristics (including their biases and personality traits as flexible/stubborn); communication functions (which simulate environmental and communal effects); and iterated learning that transfers and updates existing information from one generation to the next. 972 test cases were generated from the number of parameters included in the model and all were run.

Our results show that in some cases, dominant word order can emerge without any initial bias with different probabilities towards a particular order. We demonstrate that linguistic pressures and the rules of the language (such as fusional/logical languages) can influence a community's preference for specific word order. Contrastively, if a word order is not needed thanks to any mechanism in the language (case markings/suffixes indicate who did what to whom) and if this community is biased from the beginning towards a word order, they can accept the bias from the beginning and carry it to the future. Results also show how the first community's size and different network structures affect the dominant word order emergence and evolution speed. As the community grows, the language may accept a word order, since the differences will be difficult to keep track of and keep in mind. On the other hand, every member of the community must be connected with each other. For this, the model is tested with different network structure types (one-to-one, star, and mesh). It turns out that communities that are more connected to each other reach a regular language faster. Furthermore, communities with more flexible agents tend to change their language more readily and accept a dominant word order, while those with more stubborn individuals show little or no change in language from the beginning.

In summary, our model provides insights into the factors that contribute to the emergence and evolution of dominant word orders in natural languages. The results of this study contribute to a deeper understanding of the origins and mechanisms of word order preferences in human language while facilitating the exploration of the mechanisms underlying human learning and knowledge transfer.