

Learning and the typology of word order: a model of the Final-over-Final Condition

Overview: This paper uses computational modeling to investigate the potential influence of learnability on the typology of word order. The phenomenon of interest is the Final-over-Final Condition (FOFC: Holmberg, 2000; Sheehan et al., 2017, others), which states that a head-final phrase XP immediately dominating a head-initial phrase YP is an ill-formed syntactic structure if XP and YP are within the same extended projection. However, the FOFC does not appear to be completely categorical, as a number of apparent exceptions are attested (for example in: Bhatt and Dayal, 2007; Abels and Neeleman, 2012; Erlewine, 2017; Sheehan et al., 2017, and others). The model presented here uses a domain-general statistical learner for parameter systems, the Expectation-Driven Learner (EDL: Jarosz 2015, Nazarov & Jarosz 2017), to probe whether *FOFC languages might be difficult to learn, rather than syntactically ruled out. The EDL was tasked with learning a typology of *auxiliary*, *verb*, *object* word orders generated by a simple 4-parameter system. No parameter ruled out final-over-initial (*FOFC) languages, and no penalties targeting them were built into the EDL. Regardless, the results of the learning task demonstrate a correlation between relative learnability of word order patterns and their frequency in the typology. This suggests that a possible source of the FOFC may be the challenges final-over-initial patterns present to learning.

EDL: EDL uses a probabilistic parameteric grammar, which contains a set of probability distributions over binary parameter settings. When the EDL is presented with a training token, it samples a setting for each parameter and compares the output of the sample language to the target token. If this results in a match, parameter settings responsible for the match are rewarded. If a parameter setting contributes to a mismatch, it is penalized. Blame is assigned proportionately to each individual parameter setting’s contribution to the match/mismatch (computed using Bayes’ rule).

Learning task: The parameter space that the EDL moved through in the learning task was defined by the following:

- A) 1 parameter that enforced either harmonic head-initial or harmonic head-final word order
- B) 3 parameters that controlled movements deriving disharmonic orders from harmonic ones.
- C) The assumption that movement is always leftward (no rightward movement/specifiers)

This generated a typology of 16 parameter settings (languages) that each produced one of six surface word order patterns. Table 1 shows the patterns, the number of parameter settings that generate them, and their three tokens:

<i>Patterns:</i>	Head-Initial	Head-Final	Initial-over-final	Final-over-initial	VP-fronting	AuxP-Fronting
<i>Languages:</i>	3	5	4	2	1	1
{Aux{O}}	Aux-O	O-Aux	Aux-O	O-Aux	Aux-O	O-Aux
{V{O}}	V-O	O-V	O-V	V-O	V-O	O-V
{Aux {V {O}}}	Aux-V-O	O-V-Aux	Aux-O-V	V-O-Aux	V-Aux-O	O-Aux-V

Table 1: A simplified word order typology

The learner was trained on all languages for 80 epochs (passes through the data set) using a learning rate of 0.1. Learning was online, and averages were taken across 40 reps (i.e. individual “learners”). The learner was only exposed to strings, and had no access to any syntactic structure.

Results: Figure 1 gives the learning curves for each word order pattern. Curves are averages of each language within a pattern. As nothing in the data distinguished weakly-equivalent languages, the learner did not differentiate them.

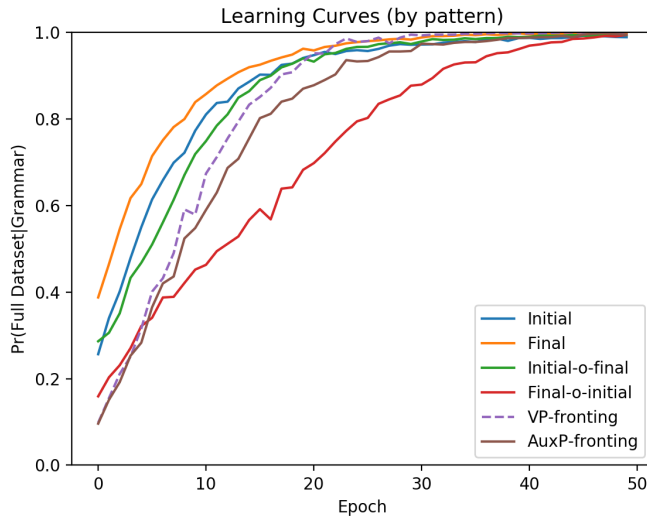


Figure 1: Learning curves for word order patterns.

The easiest word order for the EDL to learn was the harmonic head-final pattern, followed by the harmonic head-initial, and then the initial-over-final. These are the three word order patterns that are commonly attested (Greenberg, 1963). The learner struggled the most to converge on settings for the final-over-initial (*FOFC) pattern, which had the flattest learning curve. Despite increased difficulty, the *FOFC pattern was not totally unlearnable, as the EDL did converge on parameter settings when given sufficient time and exposure to data.

Discussion: The difficulty of each word order pattern was related to how strongly its tokens correlated with particular parameter settings. For example, the first parameter (controlling headedness) had the same setting across all head-final languages. Since the token *O-V-Aux* was unique to the head-final pattern, it disambiguated the setting of the first parameter for the learner. The final-over-initial pattern was difficult because it had only two languages, which had opposing settings for the first two parameters. The learner was not able to find strong evidence favouring one setting over the other, because it was pushed in two different directions. Since string-vacuous leftward movement created more weakly-equivalent languages for the initial-over-final pattern, its parameter space was smoother, making it easier to learn. Leftward movement did not create weakly-equivalent languages for the marked final-over-initial pattern. In order to create a smoother parameter space for the *FOFC order, string-vacuous rightward movement would need to be introduced to the system.

Conclusion: This learning task was conducted using a system of syntax where a) both harmonic word orders could be base-generated (contra Kayne 1994), b) disharmonic orders were derived through movement, and c) movement was always leftward. Given this system, and no further restrictions, the *FOFC pattern was more difficult for the EDL to learn than any other. The source of this difficulty was the asymmetry between leftward and rightward movement, and the effect it had on the shape of the parameter space. These results support the idea that learnability may be a source of the FOFC. It may be difficult to acquire a stable grammar for harder patterns given the limited time and data a human learner has, causing languages to shift away from them over time. If this is the case, the existence of exceptions to the FOFC are not unexpected. It may be possible to learn a *FOFC grammar, but that grammar may be unstable (i.e. not completely “set”) given the finite resources available to human learners, and may therefore be more subject to change.