Dependency length minimization: An abstract bias or an input-driven preference?

Masha Fedzechkina (University of Arizona), Charles Torres (University of Arizona), Yiyun Zhao (University of Arizona)
mfedzech@email.arizona.edu

This work addresses a long-standing question in the language sciences: whether recurrent patterns in languages across the world at least partially originate in principles of human information processing. Several studies [1, 2] have identified one abstract property that applies to all languages studied so far: although human languages differ greatly in how they order words in sentences, these superficially different orders result in short grammatical dependencies. Recent work using miniature artificial languages provided a causal link between this bias in language learners and patterns in linguistic diversity: Adult native speakers of English confronted with a novel language that had unnecessarily long grammatical dependencies systematically restructured the language to reduce dependency lengths [3].

This work leaves open an important question: Are these preferences based on general cognitive principles or on principles that are themselves learned from the statistics of the learners' native language (L1)? We tease apart these possibilities by comparing the strength of learners' preferences to reduce dependency lengths in the miniature language across English and Mandarin. These L1s are chosen because they exhibit dependency length minimization to different degrees (see Fig. 1).

Prediction: If learners whose L1 allows longer dependencies, exhibit the principle of dependency length minimization to a lesser degree in a structurally different miniature language, this behavior would suggest that miniature languages reveal an abstract principle-based L1 transfer. If, however, the degree of dependency length minimization in the miniature language is the same across learners' L1s, this would argue that this bias is rooted in pre-L1 general cognitive priors ('UG' in the broad sense).

Participants: Native speakers of English and Mandarin are recruited to participate in the experiment online via Prolific Academic, a crowdsourcing platform. The experiment is administered over the web via FindingFive, a platform for online experiment administration. Data collection has begun and will continue until the number of successful learners in each L1/miniature language reaches 20 (following [3]).

Method: The procedure and miniature languages are adopted from [3]. Participants learn a novel miniature language consisting of simple transitive sentences over 2x1h-sessions on consecutive days. Participants are exposed to a verb-final (50/50% SOV/OSV order) language (different from their verb-initial L1s). Disambiguation is achieved through obligatory case-marking on objects (never subjects). Participants first learn novel nouns (*pilika*=CHEF) and then hear sentences using these nouns along with novel verbs. During training, participants hear utterances in a novel language paired with videos of actors performing simple two-participant actions ('chef kicks referee'), where both the subject and object are either long (i.e., modified by a postpositional phrase) or short (no modification), see Fig.2. Balanced word order (SOV/OSV 50/50%) is maintained in all sentence types. Each session ends in a sentence production test: learners describe previously unseen videos in the novel language, in which constituent length is manipulated by requiring PP-modification of either the subject, object, or neither of the constituents. Constituent ordering has implications for dependency length minimization – long-before-short constituent ordering results in shorter dependencies in the verb-final miniature language.

Analysis: We calculate the average dependency length in the languages produced by individual participants and use linear regression to compare them across learners' L1's (English vs. Mandarin, sum-coded), training session (2^{nd} vs. 1^{st}, sum-coded), and their interaction.

Impact: The current study is, to our knowledge, the first attempt to understand L1 influences on processing biases in miniature artificial language learning. By teasing apart pre-L1 and L1-driven cognitive biases, we can begin to better understand how processing and L1 influences shape language learning and change and how these influences are captured in the miniature language learning paradigm.

References

1. Gildea, D. and D. Temperley, *Do grammars minimize dependency length?* Cognitive Sci, 2010. **34**(2): p. 286-310.
2. Futrell, R., K. Mahowald, and E. Gibson, *Large-scale evidence of dependency length minimization in 37 languages.* Proc Natl Acad Sci USA, 2015. **112**(33): p. 10336-10341.
3. Fedzechkina, M., B. Chu, and T.F. Jaeger, *Human information processing shapes language change.* Psychological Science, 2018.
4. Gildea, D. and T.F. Jaeger, *Human languages order information efficiently.* ArXiv e-prints, 2015.

Figures:

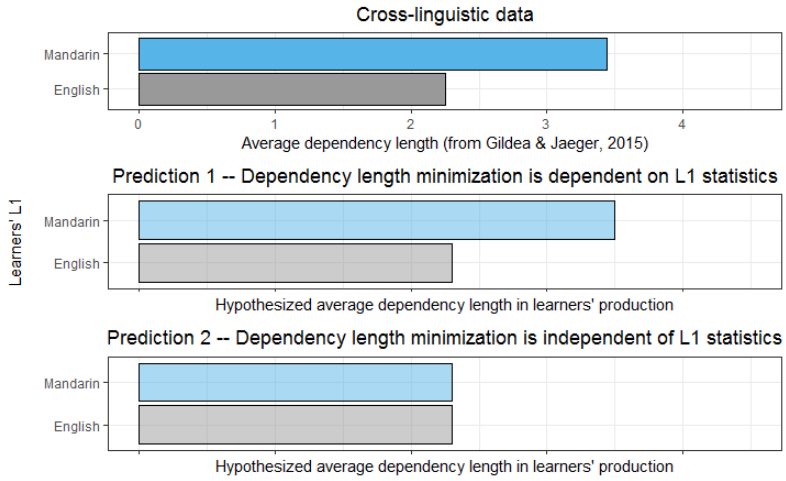

Figure 1: Illustration of the predictions for the current study. Top panel shows cross-linguistic data on dependency length minimization for the L1s used in the study (adapted from Gildea & Jaeger, 2015). Middle panel shows hypothesized dependency length reduction preferences in the miniature language if this preference is L1-dependent. Bottom panel shows hypothesized dependency length reduction preferences in the miniature language if this preference is L1-independent.

Figure 2: Sample still images of the video stimuli used in the experiment. The top panel shows the stimuli used during training: scenes with either 2 short constituents (top left) or scenes with 2 long constituents (top right). The bottom panel shows the stimuli used in testing, where only one constituent is long: either the subject (bottom left) or the object (bottom right).