Human information processing shapes language change

Short title: Information processing shapes language

Maryia Fedzechkina*[,1,2,3]
Becky Chu[4]
T. Florian Jaeger[4,5,6]

[1] Department of Linguistics, University of Arizona,
Tucson, AZ 85721 USA

[2] Graduate Interdisciplinary Program in Cognitive Science, University of Arizona,
Tucson, AZ 85721 USA

[3] Graduate Interdisciplinary Program in Second Language Acquisition and Teaching, University of Arizona,
Tucson, AZ 85721 USA

[4] Department of Brain & Cognitive Sciences, University of Rochester,
Rochester, NY 14627-0268 USA

[5] Department of Computer Science, University of Rochester,
Rochester, NY 14627-0268 USA

[6] Department of Linguistics, University of Rochester,
Rochester, NY 14627-0268 USA

*Please address correspondence to:
Maryia Fedzechkina
Department of Linguistics,
University of Arizona,
1103 E University Blvd,
Tucson, AZ 85721
Email: mfedzech@email.arizona.edu

Abstract

Human languages exhibit both striking diversity and abstract commonalities. Whether these commonalities are shaped by potentially universal principles of the human information processing has been of central interest in the language and psychological sciences. Research has identified one such abstract property in the domain of word order: although sentence word order preferences vary across languages, the superficially different orders result in short grammatical dependencies between words. As short dependencies are easier to process, these findings raise a possibility that languages are shaped by biases of human information processing. The current study directly tests the hypothesized causal link. We find that learners exposed to novel miniature artificial languages that have unnecessarily long dependencies systematically restructure the input to reduce dependency lengths rather than follow the surface preference of their native language, thus providing direct evidence for a causal link between processing preferences in individual speakers and patterns in linguistic diversity.

Keywords

Language universals; language processing; learning biases; language structure; language evolution

1. Introduction

Natural languages vary along many dimensions but this variation is not random—unrelated languages appear to share a striking number of underlying similarities. Understanding the constraints underlying these similarities has been the central question in the biological and language sciences as most theories agree that this can shed light on the mechanisms of language processing and representation in the human brain (e.g., Bates & MacWhinney, 1982; Chomsky, 1965; Christiansen & Chater, 2008; Fodor, 2001; Givón, 1991; Greenberg, 1963; Hawkins, 2014). Both constraints specific to language (Chomsky, 1965; Fodor, 2001) and constraints rooted in general principles of human information processing (Christiansen & Chater, 2008; Hawkins, 2014) have been proposed.

Focusing on the latter type, we experimentally test a hypothesized information processing constraint operating on one of the most basic and perhaps most well-studied grammatical properties of human languages—the way in which they order information in a sentence. While the order of words in a sentence varies across languages, this variability is constrained. Some word orders are cross-linguistically more frequent than others (Dryer & Haspelmath, 2011; Greenberg, 1963). Intriguingly, this cross-linguistic preference is also gradiently mirrored within languages: if a language allows several word orders, the preferred ones typically correspond to cross-linguistically common orders (Hawkins, 2014). While it has been long hypothesized that pressures associated with human information processing influence cross-linguistic word order preferences (Hawkins,

2014), the postulated causal link between the two has not yet been directly tested. We ask whether one factor that explains these preferences is a bias towards short grammatical dependencies.

Grammatical dependencies are asymmetric relations between the head (a word that licenses the presence of other words) and a dependent (a word that modifies the head). For example, in the sentence 'The boy is kicking the ball', the head (verb 'kick') forms two grammatical dependencies – one with the subject ('the boy') and one with the direct object ('the ball'). Psycholinguistic evidence shows that dependency length (i.e., the distance between the head and its dependent) affects comprehension efficiency: longer dependencies are associated with greater processing difficulty than shorter dependencies (Grodner & Gibson, 2005), an effect that is presumably due to memory retrieval (Bartek, Smith, Lewis, & Vasishth, 2011). Paralleling this comprehension advantage, language *production* also exhibits a preference for shorter dependencies. When several word order choices are available to convey the same message, speakers of verb-initial (i.e., languages that place the verb before its dependents) and verb-medial languages (i.e., languages that place the verb after the subject and before the object)like English tend to order post-verbal constituents short-before-long (Arnold, Wasow, Losongco, & Ginstrom, 2000; Wasow, 2002), while speakers of verb-final languages (i.e., languages that place the verb after the dependents) like Japanese typically prefer long-before-short ordering of pre-verbal constituents (Ros, Santesteban, Fukumura, & Laka, 2015; Yamashita & Chang, 2001). The respective verb-dependent orderings reduce the average dependency length in a sentence (Fig. 1).

While a processing advantage of shorter dependencies is well-established, its contribution to historical word order change is still under debate. Recent large-scale computational studies have provided some support for the processing account: all languages studied so far (almost 40) have average dependency lengths that are significantly shorter than would be expected by chance (Ferrer i Cancho, 2004; Futrell, Mahowald, & Gibson, 2015; Gildea & Temperley, 2010), with some languages being close to the theoretical minimum (Gildea & Temperley, 2010). While these studies provide suggestive evidence for a correlation between a preference for shorter dependencies in online processing and cross-linguistic word order constraints, they also face two critical limitations. First, typological data is sparse, making it difficult to convincingly test the validity of cross-linguistic generalizations (see debates in Croft, Bhattacharya, Kleinschmidt, Smith, & Jaeger, 2011; Dryer, 2011; Dunn, Greenhill, Levinson, & Gray, 2011). Second, and more crucially, typological data cannot directly address questions about the underlying *causes* of this hypothesized correlation. Thus, while it has been widely assumed that dependency length minimization (DLM) influences one of the fundamental abstract properties of language —grammatical constraints on word order—the causal link between the two has not so far been directly tested, leaving open the possibility that word order patterns consistent with DLM that have been observed in previous correlational studies are spurious.

Here, we test whether the cross-linguistic bias towards shorter dependencies originates in the limitations of the human processing system. Specifically, we ask whether DLM causes learners to produce languages that deviate slightly from the original input. This

would introduce small deviations into the acquired language and slightly shift the input for the next generation of learners towards a grammar with shorter dependencies. If these deviations spread through the population, small deviations can accumulate over generations, causing the language to have increasingly shorter dependencies. This would explain the tendency of natural languages to have shorter dependencies and provide evidence for processing explanations of word order constraints. If learners indeed deviate from the input towards word orders with shorter dependencies, this would support proposals that attribute certain cross-linguistic word order patterns to DLM (e.g., Hawkins, 2014). On the other hand, if learners exhibit no such preference, this would constitute a serious challenge for such accounts. By testing whether DLM causes learners to deviate from the input, we test *whether* and *how* a specific processing preference can contribute to patterns in cross-linguistic word order variation. We also, for the first time, directly assess how a preference for shorter dependencies interacts with other, already documented, learning biases [such as a bias towards simplifying the grammar, such as by fixing previously variable word order, Hudson Kam and Newport (2009)]. In this way, the present study advances our understanding of the influence of DLM and other processing preferences on word order distributions and change.

We use a miniature artificial language learning paradigm (Hudson Kam & Newport, 2009; Kirby, Tamariz, Cornish, & Smith, 2015) to directly probe the causal link between processing biases in individual learners and the DLM preference observed cross-linguistically. Miniature language learning has been successfully used to study mechanisms of first and second language acquisition (Pajak & Levy, 2014; Saffran,

Aslin, & Newport, 1996). Recent work has adapted this paradigm to explore the underlying causes of cross-linguistic patterns by creating situations of atypical input (reminiscent of situations of pidgin or language change) in the laboratory and studying how learners deviate from the atypical input they receive (Culbertson, Smolensky, & Legendre, 2012; Fedzechkina, Jaeger, & Newport, 2012; Hudson Kam & Newport, 2009; Kirby et al., 2015; Smith & Wonnacott, 2010).

Here, we present learners with input languages that have inefficient (unnecessarily long) dependencies, and test whether learners shift the language towards more efficient (shorter) dependencies.

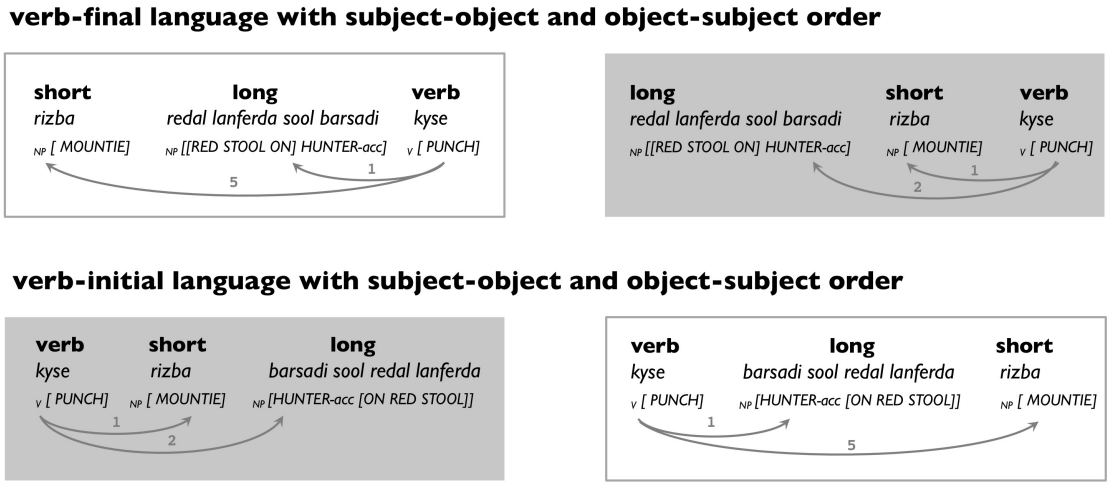

Figure 1: Comparison of dependency lengths for two possible argument orderings (subject, object) in verb-final vs. verb-initial languages. All sentences express the same meaning. Arches represent grammatical dependencies between the verb and the head of its two arguments. Numbers represent dependency lengths, measured in words. For verb-final languages (top panel), ordering dependents long-before-short leads to shorter total dependency length between the dependents and their head (the verb). For verb-initial languages (bottom panel) this relationship between the length of the dependent and its

order relative to the head is reversed: here, ordering dependents short-before-long leads to shorter overall dependency length.

2. Method

2.1. Participants

Recruiting and execution of this study was approved by the Research Subjects Review Board at the University of Rochester. Participants in the experiment were monolingual native English speakers between ages 18 and 30 recruited from the University of Rochester and the surrounding community. Each participant was exposed to only one language and received $30 for participation. To reduce the researchers' degree of freedom, recruitment continued until 20 participants successfully learned each language [following our earlier work, Fedzechkina et al. (2012); Fedzechkina, Newport, and Jaeger (2016)]. Most participants successfully learned the languages (45 participants were recruited to have 20 successful learners per language, see Section 3.1 for details).

2.2. Design and Materials

Monolingual native speakers of English learned miniature artificial languages by watching short videos describing simple transitive events performed by two human actors (e.g., 'chef punch referee') and hearing their descriptions in the novel language. Both languages had flexible word order, so that subject-object (SO) and object-subject (OS) orders occurred equally frequently in the input. Like many languages with flexible word order (Blake, 2001), our languages had consistent case-marking – a noun suffix that

disambiguated who was doing what to whom in the scene. The case-marker was always 'di' and occurred on all direct objects. The languages shared the same lexicon of 4 transitive verbs, 8 nouns (6 animate and 2 inanimate), 3 adpositions ('with', 'next-to', 'on'), and 2 color adjectives ('blue' and 'red'), see SI for more details. Both languages contained adpositional phrases (e.g., 'chef next-to blue skateboard', see Fig. 2). The order of the adposition (e.g., 'next-to') relative to its dependent ('blue skateboard') and head ('chef') followed cross-linguistically common patterns (Dryer & Haspelmath, 2011), as shown in Figure 1.

The miniature languages differed in whether they were verb-final or verb-initial. As typologically plausible, the verb-final language used pre-nominal postpositional phrases (as in Japanese or Hindi), ordering the adposition after its dependent and before its head (e.g., 'blue skateboard next-to chef'). The verb-initial language used post-nominal preposition phrases (as in English), ordering the adposition after its head and before its dependent (e.g., 'chef next-to blue skateboard').

In training, participants were exposed to sentences that either contained two 'short' constituents (i.e., both subject and object without adpositional phrase modification; 50% of training scenes) or two 'long' constituents (i.e., subject and object with adpositional phrase modification; 50% of training scenes). Sentences in which subject and object phrases differed in length were not part of the input. Word order was thus independent of phrase length in the input, and both short-short and long-long scenes occurred equally frequently with OS and SO orders. During the production test, participants described

previously unseen scenes that either contained one long constituent (either subject – in 1/3 of production scenes or object – in 1/3 of production scenes) or no modification of either constituent (1/3 of production scenes).

2.3. Procedure

The experiment was conducted in three 1-hour sessions on consecutive days with at most one day in between. Each session involved a similar combination of exposure and test blocks, with more intensive vocabulary exposure on day 1 and more intensive sentence exposure on days 2 and 3 (see Fig. 3).

Noun exposure. Participants saw pictures of characters or objects one at a time, accompanied by their names in the novel language and were instructed to repeat the names out loud to facilitate learning.

Vocabulary tests. Following noun exposure, participants completed noun comprehension and production tests. In the comprehension test, participants were shown a set of four character pictures accompanied by a name in the novel language and asked to choose the character matching the name. In the production test, participants were asked to name the character shown on the screen. Feedback on performance was provided after each trial in both tests.

Phrase exposure and tests. Participants were explicitly informed that they would learning phrases in the new language. These contained a character modified by a description (see Fig. 2 for more details). The same procedure as in vocabulary training and tests described above was used here.

Sentence exposure. Participants learned the grammar by watching short videos and hearing their descriptions in the novel language. Participants were instructed to repeat the sentences aloud to facilitate learning. On day 1, participants could replay the videos and the sound as many times as they wished; no repetitions were allowed on subsequent days.

Sentence comprehension test. Participants were presented with two side-by-side videos accompanied by auditory descriptions. The videos showed the same action and characters but the order of the actor and patient of the action reversed. Participants were asked to choose the video that matched the description. Feedback on performance was provided on each trial.

Production test. Participants were shown two novel videos side-by-side. Both videos depicted the same action and subject/object referents in switched roles (i.e., the subject referent in one video was the object referent in the other video). One of the videos was highlighted. The videos disappeared from the screen after 1200ms and were replaced by a crosshair in the center of the screen. Participants were instructed to describe the highlighted video after seeing the crosshair. A verb prompt was provided to facilitate the descriptions. No feedback on performance was provided during this test.

The use of two videos was meant to encourage participants to produce adpositional phrases—e.g., *with skateboard*—when it was present in the highlighted video. Arguably, a better way to elicit adpositional phrases may have been to present two videos that only differed in the presence of an adpositional phrase, e.g., *chef* in video 1 vs. *chef with skateboard* in video 2, rather than two videos with switched subject/object roles. However, participants overwhelmingly produced adpositional phrases as required by the scene. This is reflected in the high production accuracies reported below.

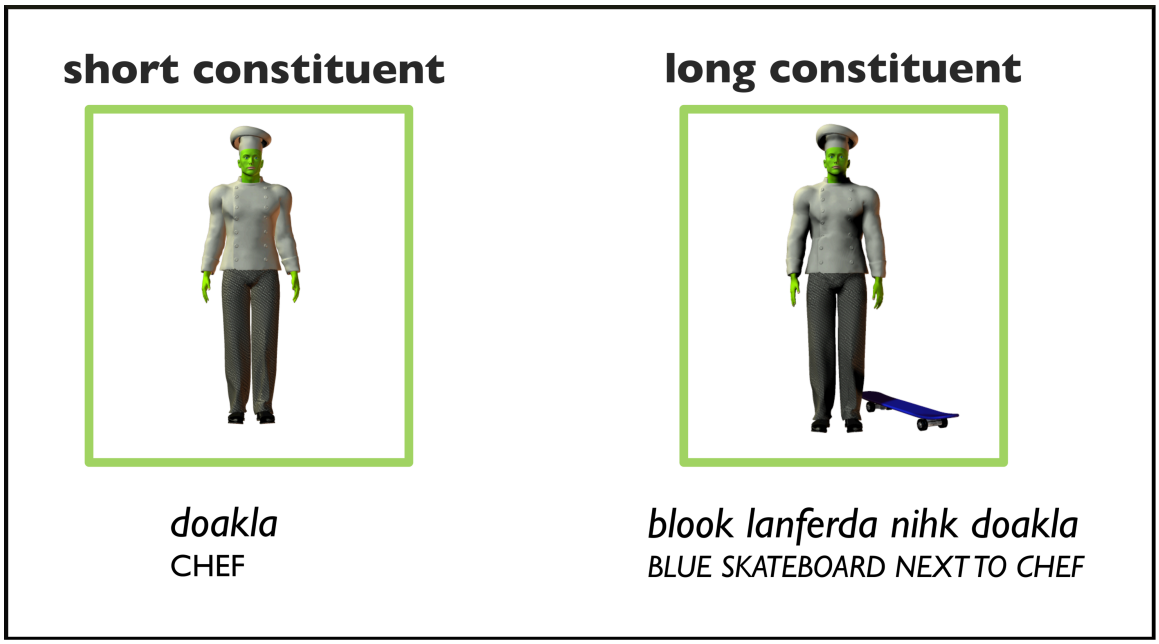

Figure 2: Illustration of constituent length manipulation in the experiment. The visual referent on the left can be described with a short phrase. The referent on the right requires a more complex long phrase. Example descriptions are shown for the verb-final miniature language (provided only auditorily in the experiment) along with their English glosses (not shown to participants).

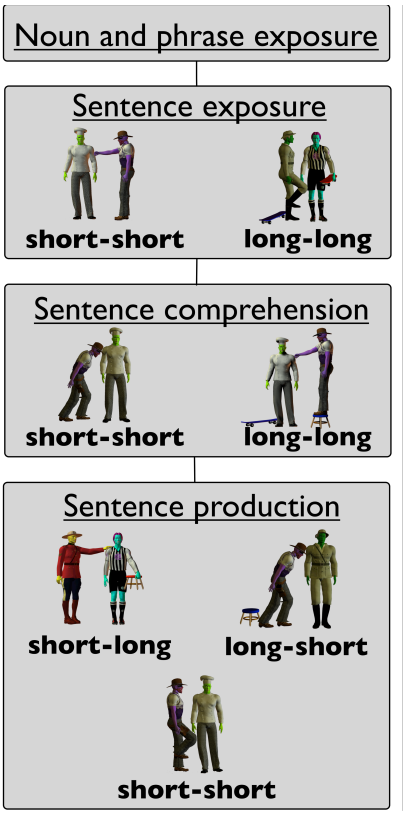

Figure 3: Each participant was exposed to a novel miniature language following the procedure shown here on each of the three separate visits to the lab (images represent still pictures of video stimuli used in the experiment). The measure of interest was the word orders learners produced in the sentence production test.

3. Results

Before turning to the predictions and central findings of our work, we describe how our data were scored and discuss learners' acquisition accuracy. We then outline the predictions of the current study and present the analyses of learners' productions.

3.1. Scoring

We first examined the accuracy of acquisition of both languages. In comprehension, we scored whether participants chose the correct video to match the sentence they heard. Since all sentences were disambiguated by case-marking, this measure allowed us to assess how well learners acquired the grammar of the novel language. Recruitment continued until the number of participants who achieved 70% accuracy on sentence comprehension tests on the final day of training reached 20 in each language. Participants who failed to pass this accuracy requirement (3 participants in the verb-final and 2 participants in the verb-initial language) were removed from further analyses. The pattern of results reported below does not depend on this exclusion.

The 40 learners submitted for further analysis achieved a high level of comprehension accuracy on the final day of training (both languages: 97% accuracy). Production performance showed a similarly high degree of accuracy, suggesting that the task was feasible (see SI for details on production scoring). In the verb-final language, participants made 8.2% lexical mistakes and 3.5% grammatical mistakes on the final day of training. In the verb-initial language, the rate of lexical mistakes was 12% and the rate of grammatical mistakes was 2.7%. All analyses reported here are based on only grammatically correct productions. We follow our previous work in not removing lexical mistakes from the analysis. The results reported below do not depend on this decision.

Given the high accuracy of acquisition of both languages, any observed word order preferences are unlikely to be due to insufficient knowledge of the lexicon and syntactic structure of the novel language.

3.2. Prediction

The central hypothesis of our study is that learners are biased towards shorter grammatical dependencies. There are two ways to assess whether such bias exists in learners' productions. We first ask whether learners order constituents within each language in a manner predicted by DLM accounts. Second, we test whether the DLM preference causes learners to deviate from the input towards significantly shorter overall dependencies, when the amount of word order flexibility in the language taken into account. These two tests complement each other: as we detail below, *whether* and *how much* participants shorten dependency length compared to the input language also depends on participants' overall word order preferences.

3.3. Relative constituent length predicts learners' word order choices in production

We begin by asking whether learners' production preferences are only affected by the surface ordering preferences in their native language or whether they are also driven by a deeper underlying principle of DLM. If learners' word order preferences are only affected by the surface order biases of their language native language, we expect learners to follow English-like short-before-long ordering (Arnold et al., 2000; Wasow, 2002). If, on the other hand, learners' word order preferences are driven by a deeper underlying principle of DLM, we expect learners to introduce a preference for shorter dependencies

in their productions. This preference should result in opposite surface orderings for the two languages: long-before-short ordering in the verb-final language and short-before-long ordering in the verb-initial language (see Fig. 1).

To assess learners' preferences in length-based ordering, we conducted a mixed effects regression analysis. We predicted learners' SO word order use from constituent length (all constituents short vs. object long, subject long vs. all other cases, Helmert coded), day of training (2 vs. 1, 3 vs. all other cases, Helmert coded), and their interactions. This analysis thus assesses learners' ordering preferences *based on the relative order of constituents within a language*, regardless of what other biases might affect *overall* word order preferences. The model contained the maximal random effects structure justified by the data based on backwards model comparison (by-subject random intercept, by-subject random slopes of day and constituent length). The same results were obtained when the maximal still converging model was used.

3.3.1. Verb-final miniature language

As expected under the DLM hypothesis, learners' word order preferences in the verb-final miniature language revealed a bias for shorter dependencies. Despite receiving an unbiased input and having the opposite short-before-long preference in their native language, learners of the verb-final language introduced a long-before-short ordering in their own productions (see Fig. 4). Across all three days of training, learners were significantly more likely to use SO order for sentences with long subject and short object phrases compared to other sentences types ($\widehat{ß}$=1.36, z=5.56, p<0.0001). Similarly,

learners were significantly more likely to use SO order for sentences in which both subject and object phrases were short, compared to sentences with short subject and long object phrases ($\hat{ß}$=0.66, z=2.54, p<0.05). There was no main effect of day of training (ps > 0.4) but day of training interacted with the effects of constituent length. On day 2, the difference in SO use between utterances with long subjects and utterances with two short constituents was significantly smaller as compared to day 1 ($\hat{ß}$=-0.56, z=-4.02, p<0.0001). The difference in SO use for sentences with two short constituents as compared to the sentences with long objects was significantly larger on day 3 compared to day 2 ($\hat{ß}$=0.18, z=2.12, p<0.05).

The analysis of simple effects revealed that learners used significantly more SO word order in utterances with long subjects compared to all other cases on all days of training (day 1: $\hat{ß}$=1.96, z=5.66, p<0.0001; day 2: $\hat{ß}$=0.84, z=3.56, p<0.0001; day 3: $\hat{ß}$=1.27, z=5.21, p<0.0001). The difference in SO word order use for sentences with long objects compared to sentences with two short constituents reached significance only on the final day of training after participants became sufficiently fluent in the novel language (day 3: $\hat{ß}$=1.02, z=3.36, p<0.0001).

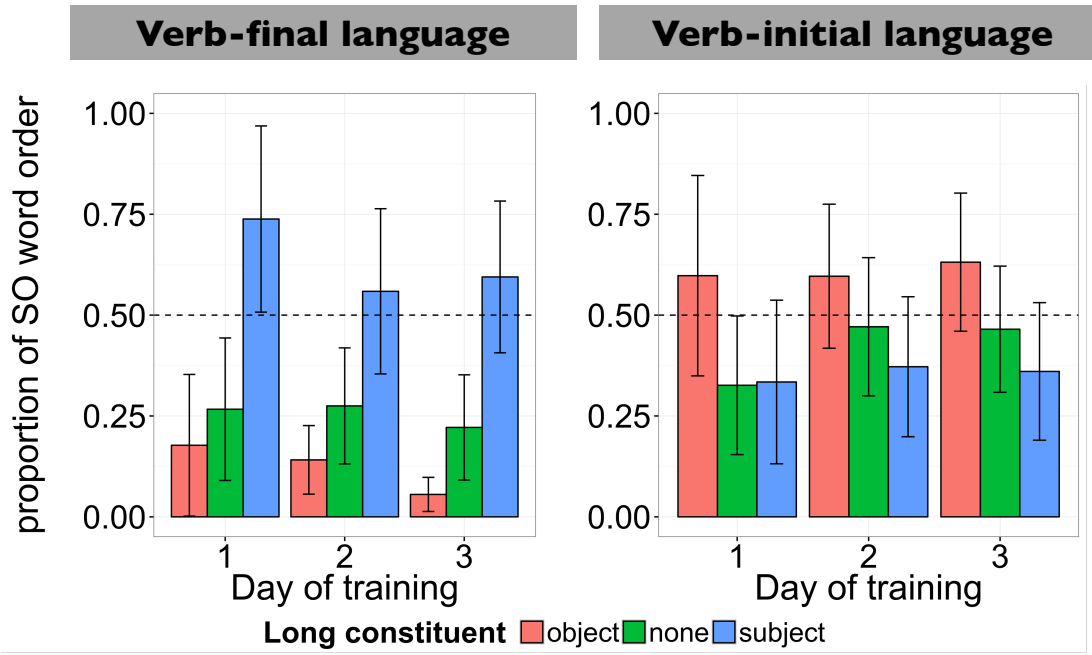

Figure 4: Subject-Object (SO) word order use in production in the verb-final (left panel) and verb-initial (right panel) languages. The dotted line indicates the input proportion of SO order (equal across all sentence types and languages). The error bars represent 95% confidence intervals.

3.3.2. Verb-initial miniature language

The verb-initial language was analyzed following the same statistical procedure and variable coding as the verb-final language.

As expected under the DLM account, learners of the verb-initial language introduced a short-before-long ordering preference in their productions—the opposite preference of that observed in the verb-final language (see Fig. 4). Across all days, learners were significantly less likely to use SO word order in sentences with long subject phrases and

short object phrases compared to all other sentence types ($\widehat{ß}$=-0.44, z=-2.5, p<0.05). Similarly, learners were significantly less likely to use SO order in sentences with short subject and object phrases compared to sentences with long object and short subject phrases ($\widehat{ß}$=-0.47, z=-2.01, p<0.05). This preference did not interact with day of training (ps > 0.2); nor there was a main effect of day of training (ps > 0.6).

Simple effects showed that the dispreference for SO word order in sentences with long subjects compared to all other sentence types was significant on day 2 ($\widehat{ß}$=-0.42, z=-2.27, p<0.05) and day 3 ($\widehat{ß}$=-0.52, z=-2.9, p<0.05) and marginally significant on day 1 ($\widehat{ß}$=-0.37, z=-1.76, p=0.08). The difference in SO use for sentences with two short constituents compared to sentences with long objects became significant with sufficient proficiency in the novel language -- on the final day of training ($\widehat{ß}$=-0.6, z=-2.47, p<0.05).

As expected under the DLM hypothesis, despite receiving an unbiased input, learners preferred opposite length-based constituent orders for verb-initial and verb-final languages, which suggests that their word order choices in production are driven by a deeper underlying preference for DLM.

The results also reveal some differences in learners' preferences across the two languages. First, the effect appears stronger in the verb-final language compared to the verb-initial language: Learners of the verb-final language introduced more pronounced

changes into the input word order than learners of the verb-initial language. Comparisons to the input discussed in the next section confirm this observation.

Second, Fig. 4 reveals that learners of the two languages differed in their overall preference for SO order. This is evident when considering only baseline (short-short) trials, for which DLM makes no ordering predictions. For baseline trials, learners of the verb-initial language matched the input on their final day of training, with 46% SO production (Wilcoxon Signed-Rank test over by-participant proportions: V= 64.5, Z=-0.39, p=0.69). Learners of the verb-final language used SO word order significantly less often than in the input (22% SO order; V= 15, Z=-3.26, $p<0.01$). These word order preferences speak against direct native language influences on learners' performance. If learners transfer surface-based ordering preferences from their native language into our experiment, we should find a preference for SO order (as in English), compared to the input. This was, however, not the case.

The bias against SO in the verb-final language is likely due to a strong preference to provide case-marking sentence-initially (case-marking occurred only on the object)—a bias we have repeatedly observed in previous work (Fedzechkina et al., 2012; Fedzechkina et al., 2016). One possible cause for this effect is a processing preference to provide informative cues sentence-initially in parsing (Hawkins, 2014; for independent evidence from artificial languages see Fedzechkina, Jaeger, & Trueswell, 2015). Given the incremental nature of sentence processing, placing a case-marked constituent sentence-initially would allow comprehenders to converge on the correct interpretation

early on and avoid costly revisions. This explanation would leave open why the bias against SO is smaller in the verb-initial language, compared to the verb-final language. One possible explanation—left to future work—is that verbs in natural languages tend to be highly informative about the correct interpretation, the information comprehenders are known to be sensitive to (Garnsey, Perlmutter, Meyers, & Lotocky, 1997), thereby reducing the perceived utility of case-marking in verb-initial languages.

Regardless of the overall difference in their preference for SO order, learners of both languages ordered longer constituents further away from the verb, as expected under the DLM hypothesis. We now ask whether the respective length-based ordering preferences in learners' productions resulted in shorter dependency lengths compared to the input.

3.4. Learners deviate from the input towards shorter dependency lengths

The analyses conducted so far show that on average word order preferences *within each language* followed the DLM prediction, when overall biases in word order use are ignored. This leaves open whether length-based orderings introduced by learners result in shorter average dependency lengths *compared to the input* when learners' overall word order preferences in the language are taken into account, as would be expected if DLM strongly affects word order preferences in learners' productions.

To address this question, we compared average per-sentence dependency length (measured in words) on the final day of training to the expected average per-sentence dependency length in the input (which did not contain length-based ordering

preferences). Baseline trials (short-short) were excluded from this analysis as DLM accounts make no prediction about word order use on these trials.

As expected under the DLM hypothesis, the output languages produced by learners had significantly shorter dependency lengths compared to what would be expected if learners reproduced the input in the verb-final (average dependency length of 3.64 on the final day of training, significantly lower than 4.5 in the input; Wilcoxon Signed-Rank test over by-participant proportions: $V= 4$, $Z=-3.06$, $p<0.001$) and verb-initial (average dependency length of 4.09 on the final day of training, significantly lower than 4.5 in the input; $V= 22$, $Z=-2.09$, $p<0.05$) languages. Overall, all but 5 out of 40 learners either match (6 learners, 15%) or reduce (28 learners, 70%) dependency length compared to the input.

Thus, the reduction in dependency length in our experiment was driven by a clear majority of the learners. As a final assessment, we quantify the degree of DLM compared to the *theoretically possible minimization*. As shown in Fig. 5, the amount of dependency length minimization that can be achieved is conditional on learners' overall word order preference, with *absolute* minimal dependency lengths only attainable if learners maintain perfectly flexible (SO vs. OS) word order. It is thus worth asking whether learners minimize dependency length *conditional on their overall SO vs. OS preference*.

20 out of 40 learners (50%) achieved the minimal theoretically possible dependency length conditional on their overall SO vs. OS preference (i.e., their productions fall on

lower gray lines in Fig. 5). This included 7 out of 9 learners who maintained perfect word order flexibility in their productions and produced output languages with dependency lengths indistinguishable from the *absolute* minimal possible dependency length. Additionally, 6 learners (15%) completely fixed word order, thereby trivially producing the minimal possible dependency length for their overall word order preference.

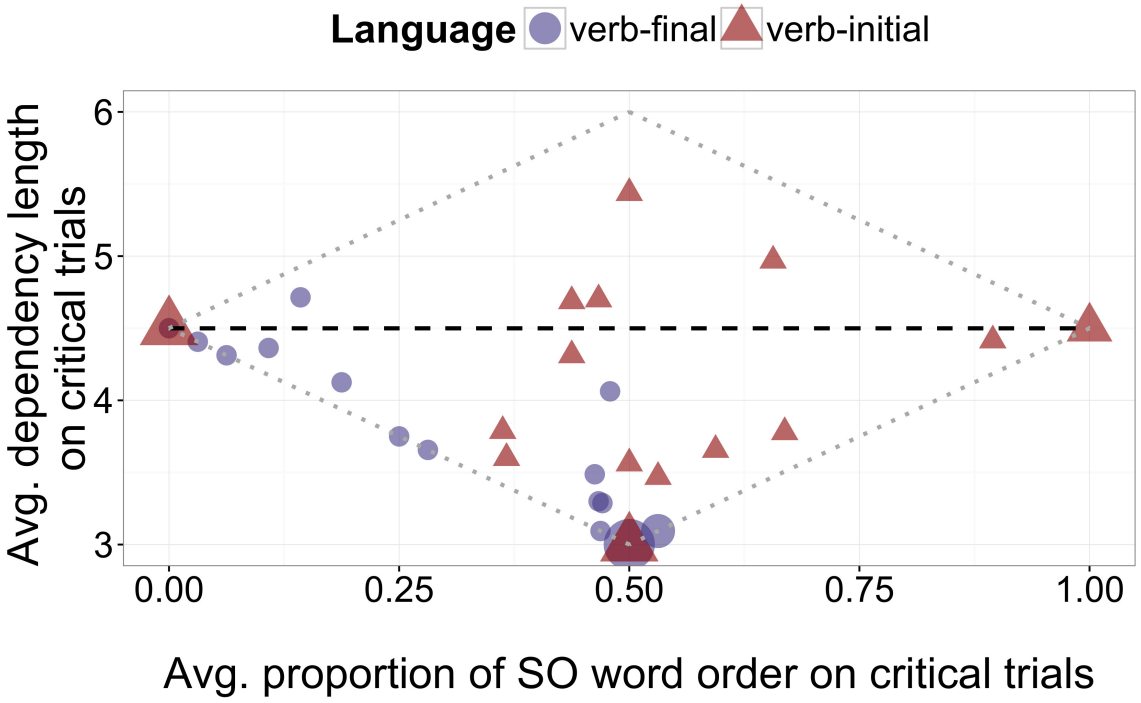

Figure 5: Average per-sentence dependency length in participants' productions on the 3rd (final) day of training plotted against the overall SO (vs. OS) ordering preference. Dependency length is measured in words. Points represent the languages produced by individual participants (point size is proportional to the number of observations per point ranging from 1 to 4). The bold dashed line shows per-sentence dependency length expected if learners exhibit no length-based ordering preferences. Participants who reduce dependency length compared to the input fall below the dashed line. The gray dotted lines outline the theoretically possible space (all learners must fall within this space). Learners who fully minimize dependency length *conditional on their overall SO preference*, fall on the lower gray lines.

Fig. 5 also reveals that learners of the verb-final language followed the DLM principle more strongly than learners of the verb-initial language. One possible explanation is that the DLM preference is enhanced when it favors a word order variant that is preferred in a language for other reasons. Recall that learners of the verb-final language produced significantly more OS order, which is consistent with a preference to provide informative cues at sentence onset. When DLM favored OS order, learners of the verb-final language followed this preference significantly more strongly than when it favored SO order (Wilcoxon Signed-Rank test over by-participant proportions: $W= 97.5$, $Z=-2.99$, $p<0.05$). Learners of the verb-initial language, who used SO and OS orders equally frequently, followed DLM equally strongly for both orders ($W= 206$, $Z=0.13$, $p=0.89$). Importantly, learners of both languages in our experiment showed a preference to reduce dependency lengths compared to the input, suggesting that the observed learning outcomes in the two languages cannot be fully explained by learners' baseline word order preferences.

Thus, both the length-based ordering preferences within each language and the reduction of dependency length compared to the input support the DLM hypothesis. As expected under the DLM accounts, learners of the verb-initial and verb-final languages introduced opposite length-based orders into their productions. Learners did so in ways that resulted in a significant reduction of the average dependency length compared to the input.

4. Discussion

The current study presents the first direct test of the hypothesized *causal link* between a processing bias for shorter grammatical dependencies and cross-linguistic word order distributions (cf. Hawkins, 2014). Our learners shared the same language background and received input languages with the same statistics but had different word order preferences depending on the verb (head) position in the language. As predicted by DLM, learners preferred short-before-long ordering in the verb-initial and long-before-short ordering in the verb-final language, which resulted in shorter dependencies in the two languages. This lends credibility to the hypothesis that the cross-linguistic preference for short dependencies originates in constraints on human information processing.

Our work adds to the debate on the role of linguistic-specific vs. domain-general in constraining word order distributions. Traditionally, grammatical constraints on word order have been explained without a reference to processing by postulating linguistic-specific generalizations such as harmony universals such as a preference to place heads either consistently before or after its dependents (Baker, 2001; Travis, 1984) or basic word order universals such as a cross-linguistic preference for SOV order (Coopmans, 1984). Later work, drawing on cross-linguistic correlational data, has proposed alternative explanations of these universal in terms of DLM—and thus, as widely assumed, in terms of human information processing (Hawkins, 2014). We find that DLM indeed influences word order distributions — at least when the input language allows two orders, learners consistently produce output languages that have shorter dependency lengths. This suggests that DLM-based explanations of harmony and basic word order

universals are plausible, making DLM a potential unifying cause behind several types of cross-linguistic word order generalizations.

Learners' preferences in our experiment are driven by an underlying DLM preference. Learners, however, do not produce languages that have optimal dependency lengths. Instead, DLM introduces small shifts into learners' productions, thus providing a seed for this cross-linguistic preference. An important open question for future research is whether these changes accumulate as the language is transmitted over generations of speakers (as assumed here), thereby causing gradual language change over historical time (cf. Christiansen & Chater, 2008; Kirby et al., 2015).

Can our findings be accounted for by learners' native language preferences? Native language transfer effects are widely attested in second language acquisition (for a review see Pajak, Fine, Kleinschmidt, & Jaeger, 2016) and thus present a serious consideration when interpreting our results. The native language of our participants (English) has an overall short-before-long preference. This could explain the result of the verb-initial miniature language, but not the inverse long-before-short preference in the verb-final language. This rules out direct surface-based transfer from English to the miniature languages as a source of the observed effects.

A related possibility is that learners transfer some form of context-specific ordering bias from English. For example, English allows topicalization ('Cheese, John already bought') and left dislocation (e.g., 'Cheese, John already bought it'). These structures realize

phrases that would otherwise occur after the verb—here the direct object—sentence-initially. There is suggestive evidence that long phrases are more likely to be topicalized or left dislocated than short phrases—a preference that is itself predicted by DLM (Snider & Zaenen, 2006). This raises the possibility that the long-before-short preference in the verb-final language is explained by a native language preference to topicalize/left dislocate long phrases. Several properties of these structures in English make this possibility rather unlikely. First, they are only licensed in specific discourse contexts (Prince, 1995), which differ from those in our experiment. Second, both structures are extremely rare in English [<0.7% reported in Gregory and Michaelis (2001)]. Low-frequency native language structures might give rise to transfer effects in miniature language studies (Goldberg, 2013). However, this would still raise the question as to why we found no evidence of a more direct transfer from English such as an overall preference for SO order.

One important question that is left open pertains to the *origin* of the DLM preference learners exhibit. Is this preference based on an innate cognitive principle or on an abstract principle acquired from the statistics of the learners' native language (Culbertson & Adger, 2014)? English exhibits DLM particularly strongly—its average dependency lengths are close to the theoretical *minimum* (Gildea & Temperley, 2010). Thus, it is possible that native speakers of English are especially attuned to DLM and are readily extending this abstract preference to the novel miniature languages. Future extensions of our work to languages with weaker DLM preferences (e.g., German or Japanese) can address this possibility. If the preference observed in our experiment is indeed learned

from the statistics of English, it raises a question of why English expresses this preference. For now, we note that DLM provides a unifying explanation for the existence of these biases both in English and in the novel miniature language. Another potentially appealing aspect of this hypothesis is that it is part of a more general proposal suggesting that the human information processing system prefers certain structures and thus can provide a parsimonious domain-general account of constraints on language structure.

Author contributions. MF and TFJ developed the study concept and contributed to the design. BC conducted the study and collected the data. MF analyzed the data with input from TFJ. All authors contributed to writing the manuscript and approved the final version of the manuscript for submission.

Acknowledgements. We thank Madeline Clark, Emily Eck, Ivy Gregory, Irene Minkina, Andy Wood, Cassandra Donatelli for help with data coding and stimuli creation.

Funding statement. This work was supported in part by NSF CAREER grant IIS-1150028 to TFJ.

Declaration of conflicting interests. The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

References:

Arnold, J. E., Wasow, T., Losongco, T., & Ginstrom, R. (2000). Heaviness vs. Newness: The effects of structural complexity and discourse status on constituent ordering. *Language, 76*(1), 28-55.

Baker, M. (2001). *The atoms of language: The mind's hidden rules of grammar*. New York: Basic Books.

Bartek, B., Smith, M., Lewis, R., & Vasishth, S. (2011). In search of on-line locality effects in sentence comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(5), 1178-1198.

Bates, E., & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L. Gleitman (Eds.), *Language Acquisition: the State of the Art* (pp. 173-218). Cambridge: Cambridge University Press.

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

Chomsky, N. (1965). *Aspects of the Theory of Syntax* (Vol. 119): The MIT press.

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behav Brain Sci, 31*(05), 489-509. doi:10.1017/S0140525X08004998

Coopmans, P. (1984). Surface word-order typology and Universal Grammar. *Language, 60*, 55-69.

Croft, W., Bhattacharya, T., Kleinschmidt, D., Smith, D. E., & Jaeger, T. F. (2011). Greenbergian universals, diachrony and statistical analyses. *Linguistic Typology, 15*(2), 433-453. doi:10.1515/lity.2011.029

Culbertson, J., & Adger, D. (2014). Language learners privilige structured meaning over surface frequency. *Proc Natl Acad Sci USA, 111*(16), 5842-5847.

Culbertson, J., Smolensky, P., & Legendre, G. (2012). Learning biases predict a word order universal. *Cognition, 122*, 306-329. doi:10.1016/j.cognition.2011.10.017

Dryer, M. (2011). Evidence for word order correlations: A response to Dunn, Greenhill, Levinson and Gray's paper in Nature. *Linguistic Typology, 15*, 335-380.

Dryer, M., & Haspelmath, M. (Eds.). (2011). *The world atlas of language structures online*. Munich: Max Planck Digital Library.

Dunn, M., Greenhill, S. J., Levinson, S. C., & Gray, R. D. (2011). Evolved structure of language shows lineage-specific trends in word-order universals. *Nature, 473*(7345), 79-82. doi:10.1038/nature09923

Fedzechkina, M., Jaeger, T., & Newport, E. (2012). Language learners restructure their input to facilitate efficient communication. *Proc Natl Acad Sci USA, 109*(44), 17897-17902. doi:10.1073/pnas.1215776109

Fedzechkina, M., Jaeger, T., & Trueswell, J. (2015). Production is biased to provide informative cues early: Evidence from miniature artificial languages. In D. Noelle, A. Waelaumont, J. Yoshimi, T. Matlock, C. Jennings, & P. Maglio (Eds.), *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 674-679). Austin, TX: Cognitive Science Society.

Fedzechkina, M., Newport, E., & Jaeger, T. (2016). Balancing effort and information transmission during language acquisition: Evidence from word order and case-marking. *Cognitive Science*. doi:doi: 10.1111/cogs.12346

Ferrer i Cancho, R. (2004). Euclidean distance between syntactically linked words. *Physical Review E, 70*, 056135.

Fodor, J. D. (2001). Setting syntactic parameters. *The Handbook of contemporary syntactic theory*, 730-767.

Futrell, R., Mahowald, K., & Gibson, E. (2015). Large-scale evidence of dependency length minimization in 37 languages. *Proc Natl Acad Sci USA, 112*(33), 10336-10341.

Garnsey, S. M., Perlmutter, N. J., Meyers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language, 37*, 58–93.

Gildea, D., & Temperley, D. (2010). Do grammars minimize dependency length? *Cognitive Sci, 34*(2), 286-310.

Givón, T. (1991). Markedness in grammar: Distributional, communicative and cognitive correlates of syntactic structure. *Stud Lang, 15*(2), 335-370. doi:10.1075/sl.15.2.05giv

Goldberg, A. (2013). Substantive learning bias or an effect of similarity? Comment on Culbertson, Smolensky, Legendre (2012). *Cognition, 127*(3), 420-426.

Greenberg, J. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. Greenberg (Ed.), *Universals of human language* (pp. 73-113): MIT Press.

Gregory, M., & Michaelis, L. (2001). Topicalization and left-dislocation: a functional opposition revisited. *Journal of pragmatics, 33*(11), 1665-1706.

Grodner, D., & Gibson, E. (2005). Consequences of the serial nature of linguistic input. *Cognitive Science, 29*(2), 261–290.

Hawkins, J. A. (2014). *Cross-linguistic variation and efficiency*. Oxford: OUP.

Hudson Kam, C., & Newport, E. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychol, 59*(1), 30-66.

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition, 141*, 87-102.

Pajak, B., Fine, A. B., Kleinschmidt, D., & Jaeger, T. (2016). Learning additional languages as hierarchical inference: insigts from L1 processing. *Language Learning*.

Pajak, B., & Levy, R. (2014). The role of abstraction in non-native speech perception. *Journal of Phonetics, 46*, 147-160.

Prince, E. F. (1995). On the limits of syntax, with reference to left-dislocation and topicalization. In P. W. Culicover & L. McNally (Eds.), *The limits of syntax* (pp. 281-302). San Diego, CA: Academic Press.

Ros, I., Santesteban, M., Fukumura, K., & Laka, I. (2015). Aiming at shorter dependencies: The role of agreement morphology. *Language, Cognition, and Neuroscience, 30*(9), 1156-1174.

Saffran, J., Aslin, R., & Newport, E. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926-1928.

Smith, K., & Wonnacott, E. (2010). Eliminating unpredictable variation through iterated learning. *Cognition, 116*(3), 444-449. doi:10.1016/j.cognition.2010.06.004

Snider, N., & Zaenen, A. (2006). Animacy and syntactic structure: Fronted NPs in English. In M. Butt, M. Dalrymple, & T. H. King (Eds.), *Intelligent Linguistic Architectures: Variations on Themes by Ronald M. Kaplan*. Stanford: CSLI Publications.

Travis, L. (1984). *Parameters and Effects of Word Order Variation.* (PhD dissertation), MIT.

Wasow, T. (2002). *Postverbal Behavior*. Stanford: CSLI Publications.

Yamashita, H., & Chang, F. (2001). Long before short preference in the production of a head-final language. *Cognition, 81*(2), B45-B55.