Is the strength of regularisation behaviour uniform across linguistic levels?

Carmen Saldana (C.C.Saldana@sms.ed.ac.uk), Kenny Smith, Simon Kirby & Jennifer Culbertson
Centre for Language Evolution, School of Philosophy, Psychology and Language Sciences,
The University of Edinburgh, Dugald Stewart Building, 3 Charles Street, Edinburgh, EH8 9AD, UK

Abstract

Human languages contain very little unconditioned variation. In contexts where language learners are exposed to input that contains inconsistencies, they tend to regularise it, either by eliminating competing variants, or conditioning variant use on the context. In the present study we compare regularisation behaviour across linguistic levels, looking at how adult learners respond to variability in morphology and word order. Our results suggest similar strengths in regularisation between linguistic levels given input languages whose complexity is comparable.

Keywords: artificial language learning; statistical learning; regularisation; variation; complexity; morphology; word order

Introduction

While languages exhibit variation at all linguistic levels, in the form of paraphrases, synonyms, allomorphs and allophones, that variation tends to be predictable: the choice of variant is (at least partially) conditioned by some aspect of the social or linguistic context. Occasionally, language learners are exposed to input that involves inconsistencies, for instance, when new variants are introduced into an established system, or when conventions are still not established, as in emerging languages (Senghas & Coppola, 2001; Siegel, 2004). Learners under those circumstances tend to reduce or remove such inconsistencies, i.e. they regularise their input. This can be achieved either by removing competing variants, or conditioning variant choice on the context (Ferdinand, Kirby, & Smith, 2017).

Regularisation has been documented extensively across linguistic levels (i.e. phonology, morphology, syntax and the lexicon) in natural language; e.g. in language acquisition, language change, and in emerging languages (Senghas & Coppola, 2001; Siegel, 2004; van Trijp, 2013). Experimental studies involving artificial language learning and statistical learning techniques report regularisation behaviour during the learning and production of probabilistic unconditioned variation in different linguistic units, across different linguistic levels (Culbertson, Smolensky, & Legendre, 2012; Fehér, Wonnacott, & Smith, 2016; Hudson Kam & Newport, 2005, 2009; Wonnacott & Newport, 2005). Nevertheless, it still remains an open question whether regularisation behaviour applies with uniform strength across linguistic levels and to what extent level-specific biases interact with regularisation during language learning and use.

Level-specific effects in regularisation behaviour

Research in second language acquisition and pidgin and creole studies has highlighted different developmental paths for morphology and syntax cross-linguistically (Good, 2015; Slabakova, 2013). Studies in pidginisation suggest that, in periods when pidgins are highly inconsistent, linguistic levels might behave differently: morphologically complex traits such as inflectional morphology seem to be highly simplified whilst syntactic traits such as word order tend to reproduce the input complexity more closely (Good, 2015; Siegel, 2004). Good (2015) argues that this asymmetry is given by a break in transmission from source languages for morphological traits, which are only successfully transmitted if an entire contrasting paradigm is available to the learner, which is not the case in periods of linguistic instability. However, word order variation can be contrastive as well (e.g. S-Aux inversion to distinguish illocutionary forces). Alternatively, a more parsimonious hypothesis we could entertain is that a general tendency for pidgins to comprise highly simplified morphological traits and more conservative word order is rooted in the differing complexity of these traits in the source languages; Hudson Kam and Newport (2009) show that learners are more likely to regularise complex systems of variation.

Recent experimental studies have separately explored the effect of learning biases on typological asymmetries found in morphology and word order respectively. In morphology for example, St Clair, Monaghan, and Ramscar (2009) provide evidence of a preference for suffixing over prefixing, mirroring the cross-linguistic preference for suffixing. In word order, Culbertson et al. (2012) show that learners prefer consistent harmonic word order patterns (i.e. all modifiers either pre-nominal or post-nominal), also found more commonly in the world's languages. Moreover, Culbertson et al. (2012) show that this bias leads to different regularisation behaviour for different word order patterns. Nevertheless, no study has hitherto tried to systematically compare regularisation behaviour *across* linguistic levels. Uncovering differences in regularisation behaviour across linguistic levels could shed light on the intriguing asymmetry found in pidgin languages: morphological paradigms seem to be highly simplified whilst input complexity is more closely reproduced in word order.

In the present study we combine artificial language learning and statistical learning techniques to systematically compare the strength of regularisation of inflectional morphology and word order, controlling for asymmetries in the complexity and variability of the input languages.

Experiment 1

We utilise the methodology developed in Culbertson et al. (2012); Hudson Kam and Newport (2005). Adult learners are exposed to a miniature artificial language featuring an inconsistent mixture of synonymous variants. We are interested in how learners restructure the probabilistic unconditioned variation in the input languages, and to what extent that restructuring is comparable across linguistic levels (specifically, morphology and word order).

Method

Participants Fifty-six native-English speakers (aged between 18 and 41, *mean* = 23.2) were recruited from the University of Edinburgh's Careers Service database of vacancies. Each was compensated £6. Twenty-six participants were assigned to the Morphology condition, and 26 to the Word Order condition; the data from a further 4 participants (all in the morphology condition) were excluded as they either failed to learn the noun lexicon or failed to learn the associations between phrases and pictures.

Input languages We designed two novel languages which contained probabilistic unconditioned variation either in morphology or word order. Their respective probabilistic grammars are shown in Table 1. Both languages were used to describe simple pictures featuring one of two objects. Each object appeared either singly or in a pair; and could appear either in greyscale or coloured in blue. Descriptions were noun phrases composed of a Noun plus a Num(eral) and/or Adj(ective) modifier, which were presented orthographically and aurally to participants during the experiment.

All lexical items were 5 graphemes/phonemes long and had a neighbourhood density of 0 in the English lexicon. Nouns and modifiers differed in their syllabic structure; while all were bisyllabic, nouns (i.e. "mokte" and "jelpa") conformed to a CVC.CV pattern, and modifiers to CV.CCV (based on English phonotactics and the Maximal Onset Principle).

Procedure Participants worked through a six-stage training and testing regime.

Stage 1, noun familiarisation Participants were trained on the two bare nouns that corresponded to pictures of the two different objects in the artificial language. During this phase, participants underwent a block of training consisting of 6 exposure trials and 4 picture-selection comprehension trials (in that order) —each noun-picture pair appeared 5 times (order randomised). Common to all training blocks to follow, on each exposure trial participants were presented with a picture (in this block always of a single object in grey-scale) and a corresponding description in the language (in this block, a bare noun), displayed both visually and aurally. On comprehension trials, participants were asked to select a picture out of an array of four (in this stage, the two objects seen during training plus two distractors) that corresponded to the displayed description in the alien language, and received feedback on their accuracy.

Stage 2, one-modifier training In Stage 2 participants were trained on one-modifier NPs, i.e. a Noun plus either Num or Adj only. Pictures contained any of the two objects presented either in blue and singly (Adj only) or in greyscale and in pairs (Num only). For each picture, a variant was selected randomly from the grammar assigned to the participant. Both grammars contained majority variants with an empirical probability of $P = 0.6$, and minority variants with $P = 0.4$, as shown in Table 1. This phase comprised 40 trials in total, divided in 2 blocks of 20 trials; each block consisted of 15 exposure trials followed by 5 picture-selection trials. Participants saw each of the four different one-modifier pictures 5 times per block (order randomised).

Table 1: Probabilistic input languages in the Morphology and Word order conditions. Languages contain probabilistic unconditioned variation in inflectional morphology or word order respectively. All morphological variation resides in the suffixation of the modifiers. All word order variants conform to constituent structure [Num [Adj N]]. There are three types of NPs: Num Only (single Num modifier) refer to objects in pairs and in grey-scale, Adj Only (single Adj modifier) refer to a single object coloured in blue, and two-Mod(ifier) NPs (with both Num and Adj modifiers) correspond to objects in pairs coloured in blue. Languages include two different nouns (each corresponding to a different object) and thus comprise a total of 16 NPs (8 per noun) that correspond to a total of 6 pictures (1 per NP type, 3 per object).

NP TYPE	MORPHOLOGY CONDITION	WORD ORDER CONDITION
NUM ONLY	0.6 $NP \rightarrow$ **N nefri**	0.6 $NP \rightarrow$ **N nefri**
	0.4 $NP \rightarrow$ N nezno	0.4 $NP \rightarrow$ nefri N
ADJ ONLY	0.6 $NP \rightarrow$ **N kogla**	0.6 $NP \rightarrow$ **N kogla**
	0.4 $NP \rightarrow$ N kospu	0.4 $NP \rightarrow$ kogla N
TWO MOD	0.6 $NP \rightarrow$ **N kogla nefri**	0.6 $NP \rightarrow$ **N kogla nefri**
	$0.1\bar{3}$ $NP \rightarrow$ N kogla nezno	$0.1\bar{3}$ $NP \rightarrow$ nefri kogla N
	$0.1\bar{3}$ $NP \rightarrow$ N kospu nefri	$0.1\bar{3}$ $NP \rightarrow$ nefri N kogla
	$0.1\bar{3}$ $NP \rightarrow$ N kospu nezno	$0.1\bar{3}$ $NP \rightarrow$ kogla N nefri

Stage 3, one-modifier testing Stage 3 of the experiment tested the participants' knowledge of the language. They saw the same pictures used in Stage 2 without accompanying text or audio and were asked to type in an appropiate description. They had to describe 20 pictures in total; each of the four different one-modifier pictures was presented 5 times in random order.

Stage 4, full training In Stage 4 participants were trained on a mix of one-modifier (a noun plus Adj or Num) and two-modifier NPs (a noun plus both Num and Adj). Two-modifier NPs were used to describe pairs of blue objects. For one-modifier phrases, variants were chosen in the same way as in Stage 2. For two-modifier phrases, variants were also selected randomly from the grammars assigned, with empirical probabilities of $P = 0.6$ and $P = 0.1\bar{3}$ for the majority and the three minority variants respectively (see Table 1). This stage comprised 100 trials (20 Num Only, 20 Adj Only and 60 two-Mod), divided into 4 block of 25 (15 exposure training trials followed by 10 picture-selection trials). Participants saw each of the four one-modifier pictures 10 times, and each of the two two-modifier pictures 30 times.

Stage 5, full testing Stage 5 tested participants' knowledge of the whole language. They saw all pictures they had been trained on and were asked to type in appropriate descriptions. They had to describe 52 pictures in total: 10 Adj Only (5 per object), 10 Num Only (5 per object), 30 two-modifier (15 per object), and additionally, 2 pictures of bare objects by themselves and in grey-scale (1 per object).

Results

Output variability Figure 1 shows the entropy of participants' production systems for both the Morphology and Word Order conditions. Analyses are run on Stage 5's testing exclusively, i.e. participants' final production sets. Words in the productions were corrected for typos (and only typos). Shannon entropy measures how variable participants' productions are; the higher the scores, the more variable and the lower the scores, the more regular. The Shannon entropy (H) of phrase use for participant is given by

$$H(X) = -\sum_{i=1}^{n} P(x_i) log_2 P(x_i) \tag{1}$$

where the sum is over the different variants, and $P(x_i)$ is the empirical probability of variant x_i in the set of a participant's productions, X. We treated the two nouns for the different objects as the same variant when we calculated the entropy of the phrase variants such that no variability is introduced by the correct use of the different nouns. Entropy lower- and upper- bounds will vary depending on the number of required and possible variants as well as on the number of production trials. The most regular expressive language contains only one-to-one picture-phrase mappings and therefore only three different variants, one Num Only (e.g. *N nefri*), one Adj Only (e.g. *N kogla*) and one two-modifier (e.g. *N kogla nefri*). The final production phase consisted of 50 trials (excluding the two bare noun trials), divided up into 20 one-modifier trials (half Num Only and half Adj Only) and 30 two-modifier trials: the entropy lower bound for the language *overall* is thus 1.37 bits, and 0 bits for each of the NP types.

Figure 1 shows the entropy scores for the set of all participants' productions (i.e. the overall language), as well as those for the production sets for specific NP types in isolation: one-modifier Num (Num Only), one-modifier Adj (Adj Only), and two-modifier (two-Mod) NPs. Entropy lower bounds and input entropies are represented as solid and dotted vertical lines respectively. A visual inspection of the Morphology and Word Order conditions in Figure 1 suggests that in many cases participants failed to reproduce the full variability of the input languages; entropy scores are generally lower.

We used the *stats* and *lme4* packages developed in R (Bates, Mächler, Bolker, & Walker, 2015; R Core Team, 2015) to run a linear mixed effects regression model (which we will call Model 1) to explore the effect of condition on regularisation behaviour (dependent variable: entropy). As fixed effects we entered Condition (two levels: Morphology as reference, and Word Order), NP Type (reverse Helmert coded with the 3 ordered levels: Num Only, Adj Only and two-Mod) and System (two levels: Input as reference, and Output). We also entered all interactions between fixed effects. As random effects, we included intercepts for Subject as well as by-Subject slopes for the effects of NP Type and System type. P-values were obtained through the *lmerTest* package (Kuznetsova, Bruun Brockhoff, & Haubo Bojesen Christensen, 2015). Results show a significant effect of System ($\beta = -0.346$, $SE = 0.085$, $p < .001$), suggesting that participants did indeed regularise their input in their output productions. We also found a significant interaction between System and Condition ($\beta = -0.284$, $SE = 0.119$, $p = .021$), suggesting that participants regularised their input significantly more in the Word Order condition. Results show the expected effect of higher input entropies in two-Mod NPs ($\beta = 0.21$, $SE = 0.024$, $p =< .001$), and no significant interactions between NP Type and System (largest: $\beta = 0.027$, $SE = 0.028$, $p = .324$) or between NP Type, System and Condition (largest: $\beta = -0.041$, $SE = 0.039$, $p = .299$). These results suggest that participants regularised their input systems across conditions and NP types, and that participants in the Word Order condition regularised them more than those in the Morphology condition.

Table 2: Central tendencies of the proportion of majority input variants in production by condition and NP type. From left to right, the mean, median and mode(s).

		Proportion Majority Input Variant in Production		
		mean	*median*	*mode(s)*
Morphology	Num Only	0.704	0.8	0.919
	Adj Only	0.669	0.7	0.916
	two-Mod	0.609	0.65	0.843
Word Order	Num Only	0.580	0.65	0.094 & 0.96
	Adj Only	0.585	0.7	0.104 & 0.947
	two-Mod	0.442	0.33	0.089 & 0.92

Variant production Table 2 provides the central tendencies for proportion use of the majority input variant for each NP type. We observe that all distributions in the Word Order condition are bimodal, with modes of the distributions of majority variant use at $P \leq 0.1$ and $P > 0.9$ across NP types, suggesting two opposite trends amongst participants: one towards the over-production of the majority input word order variants and another, towards their under-production.

Participants under-producing the majority word order variant in one-modifier NPs are necessarily producing modifiers pre-nominally. Figure 2 shows the overall proportions of the variants produced for two-Mod NPs by all participants. The input proportions are represented by the yellow vertical lines. The word order produced the most is the majority input variant N Adj Num. Although the three remaining input variants (below the grey solid line division) were equally frequent in the input language, the Num Adj N word order is overall

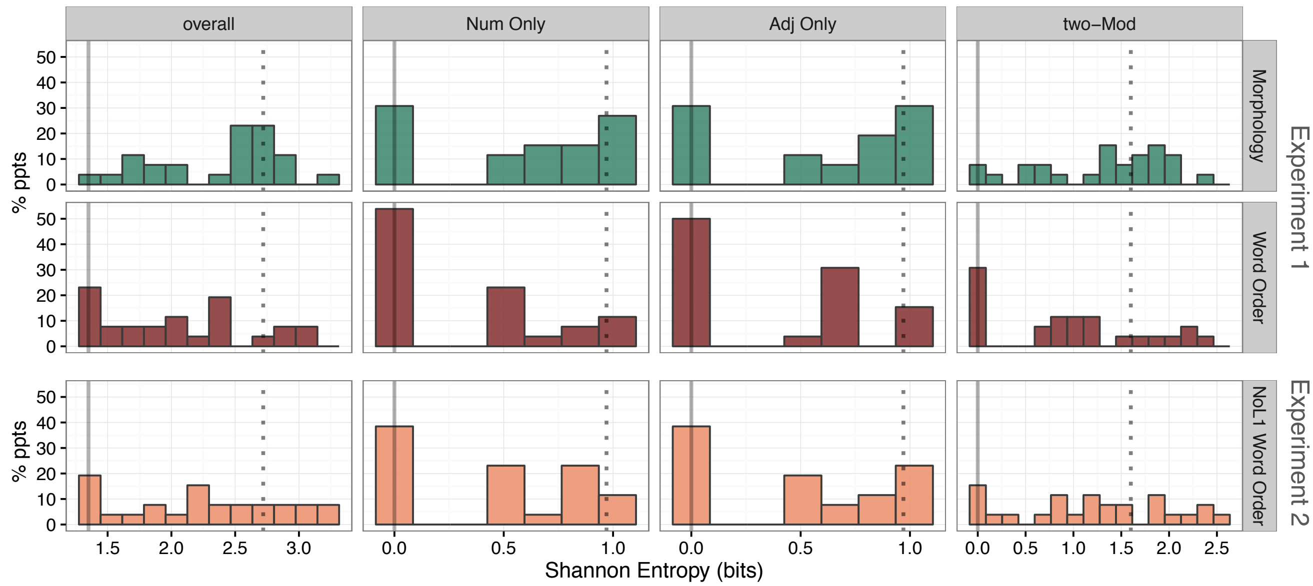

Figure 1: Entropy scores of participants' production systems. From top to bottom, scores for the Morphology (green) and Word Order (red) conditions in Experiment 1 and for the NoL1 Word Order condition (orange) in Experiment 2. From left to right, entropies of participants' full production sets as well as entropies by NP type: one-modifier Num (Num Only), one-modifier Adj (Adj Only), and two-modifier (two-Mod) NPs. Input entropy scores are indicated by dashed vertical lines. Minimum entropy scores are indicated by solid vertical lines. Minimum entropy is always 0 for each NP type in isolation but 1.37 for the overall system as it necessitates a minimum of 3 variants, one per NP type.

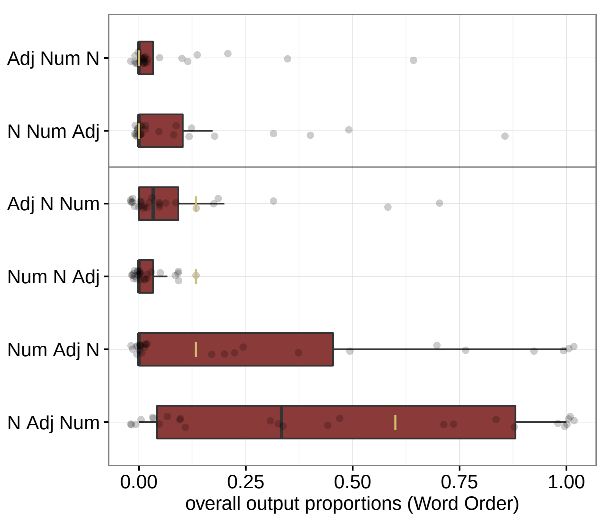

Figure 2: Box plot displaying the output proportions of two-modifier variants in the Word Order condition with individual participants' data points overlaid. Seen (bottom) and unseen (top) variants during training are divided by a solid grey line. Vertical yellow lines indicate input proportions.

more frequently used (although only by a minority as indicated by the median value 0). Only 30% of participants produced systems with both harmonic variants (Num Adj N and N Adj Num) —and only 19% produced both variants more than once, suggesting that although both harmonic orders are preferred overall, they do not generally coexist within the productions of a single participant.

We ran a logistic regression model, which we will call Model 2, to explore the average difference between the proportions of Num Adj N variants in input and output linguistic systems. We entered System (two levels: Input as reference, and Output) as the only fixed effect. Random intercepts for Subject as well as by-Subject random slopes for the effect of System were also included. Results show that the Num Adj N variant is produced significantly less in output languages than in the input language ($\beta = -7.641$, $SE = 1.943$, $p < .001$). Only a minority of participants overproduced this variant, the majority of participants were in fact under-producing it. On top of the observed preference for harmonic order, these results confirm a tendency to avoid systems with two opposite N-peripheral variants, i.e. N Adj Num and Num Adj N.

Discussion of Experiment 1

Our results provide evidence that learners regularise probabilistic unconditioned variation in both morphology and word order. Regularisation behaviour is in line with an overarching simplicity bias argued to be at play in language learning and use (Culbertson & Kirby, 2016). Though the input languages were similar in terms of overall system complexity, regularisation behaviour was slightly stronger in the Word Order condition than in the Morphology condition. A close analysis of the variant usage in the Word Order condition suggests that this difference is driven by a bias in favour of harmonic N Adj Num and Num Adj N variants but against their coexistence within a system. This bias could be the result of L1 transfer; participants may have overproduced the Num Adj

Table 3: Probabilistic input language in the NoL1 Word order condition in contrast to the Word Order condition in Experiment 1. Changes in the variant set are indicated with boxes.

NP TYPE	WORD ORDER	NOL1 WORD ORDER
NUM ONLY	$0.6\ NP \rightarrow$ **N nefri**	$0.6\ NP \rightarrow$ **N nefri**
	$0.4\ NP \rightarrow$ nefri N	$0.4\ NP \rightarrow$ nefri N
ADJ ONLY	$0.6\ NP \rightarrow$ **N kogla**	$0.6\ NP \rightarrow$ **N kogla**
	$0.4\ NP \rightarrow$ kogla N	$0.4\ NP \rightarrow$ kogla N
TWO MOD	$0.6\ NP \rightarrow$ **N kogla nefri**	$0.6\ NP \rightarrow$ **N kogla nefri**
	$\boxed{0.1\bar{3}\ NP \rightarrow \text{nefri kogla N}}$	$\boxed{0.1\bar{3}\ NP \rightarrow \text{N nefri kogla}}$
	$0.1\bar{3}\ NP \rightarrow$ nefri N kogla	$0.1\bar{3}\ NP \rightarrow$ nefri N kogla
	$0.1\bar{3}\ NP \rightarrow$ kogla N nefri	$0.1\bar{3}\ NP \rightarrow$ kogla N nefri

N word order because it is the most common order in their L1 grammar. To minimise the possible effects of this level-specific word order bias, Experiment 2 investigated learning in a second word order condition, removing the English-like two-modifier harmonic pattern from the input.

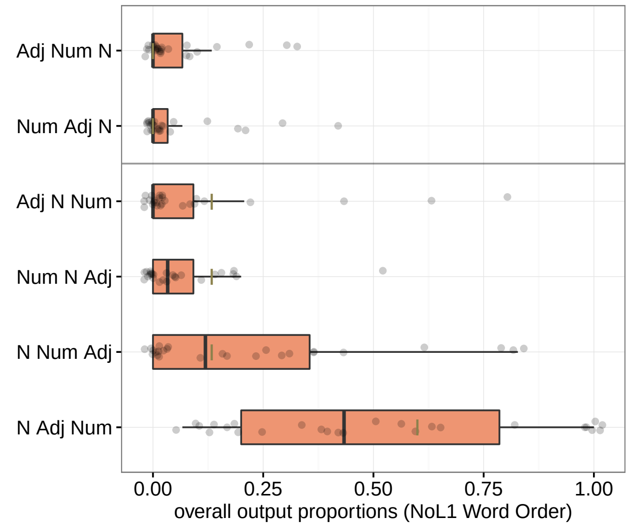

Figure 3: Box plot displaying the output proportions of two-modifier variants in the NoL1 Word Order condition with individual participants' data points overlaid. Divided by a solid grey line, seen (bottom) and unseen (top) variants during training. Vertical light brown lines indicate input proportions.

Experiment 2

Experiment 2 follows the same design as the Word Order condition described in Experiment 1, with one difference: the set of two-modifier NP input variants. As illustrated in Table 3, we replaced the Num Adj N variant with the N Num Adj pattern, maintaining the number of harmonic word orders (two, i.e. N Adj Num and N Num Adj) but eliminating the L1 variant and the presence of opposite N-peripheral patterns. For ease of reference, we call Experiment 2 the NoL1 Word Order condition. We expect the change in the input language to mitigate the effect of L1 transfer and to increase the coexistence of both harmonic patterns.

Participants Twenty-eight native-English speakers (aged between 18 and 35, *mean* = 24.8) were recruited via the University of Edinburgh's Careers Service advertisement database. Participants received £6. Only the data from 26 participants were fit for analysis as two participants either failed to learn the noun lexicon or failed to learn the associations between phrases and pictures.

Results

Entropy scores obtained in the NoL1 Word Order condition are shown in Figure 1 (coloured in orange). We ran a linear mixed effects model as in Experiment 1 to explore the effect of condition on regularisation behaviour (dependent variable: entropy), including the conditions in Experiment 1 plus NoL1 Word Order. The mixed-effects structure was the same as in the Model 1 but with reverse Helmert coding of Condition such that NoL1 Word Order was directly compared to the Morphology condition from Experiment 1, and the Word Order condition was compared to the average of the Morphology and NoL1 Word Order conditions. Results show a significant effect of System ($\beta = -0.483$, $SE = 0.051$, $p < .001$) and a significant interaction between Word Order and System ($\beta = -0.073$, $SE = 0.036$, $p = .046$), ratifying the results in Model 1. However, we did not find a significant interaction between NoL1 Word Order and System ($\beta = -0.063$, $SE = 0.063$, $p = .317$), suggesting that participants in the Morphology and the NoL1 Word Order conditions regularised their input to similar degrees, and on average they regularised it less than participants in the Word Order condition in Experiment 1. As in Model 1, we did not find significant interactions between NP Type and System (largest: $\beta = 0.016$, $SE = 0.015$, $p = .288$) or between NP Type, System and Condition (largest: $\beta = -0.015$, $SE = 0.011$, $p = .168$). These results suggest that participants regularised their input systems across conditions and NP types, and that whilst participants in the Word Order condition regularised more than those in the Morphology condition, participants in the Morphology and the NoL1 Word Order conditions regularised their input to similar degrees. Excluding the Num Adj N variant in the input language thus eliminated the difference between levels. In other words, participants do not regularise probabilistic unconditioned variation in word order more than in morphology.

Figure 3 shows the overall proportions of the variants produced for two-Mod NPs in the NoL1 Word Order condition. We observe that the most produced word order is the majority input variant N Adj Num, and that the harmonic N Num

Adj word order is overall more frequent than any other minority input variant. Unlike in the Word Order condition where systems with both Num Adj N and N Adj Num patterns were not common, 65% of participants produced systems with both N Adj Num and N Num Adj harmonic variants in the NoL1 Word Order condition. We ran a logistic regression model to test the difference between the proportions of N Num Adj variants in input and output linguistic systems across participants. We used the same mixed-effects structure as in Model 2. Results suggest that the proportion of N Num Adj variants in the output languages is not significantly different from the input proportion across participants ($\beta = -0.594$, $SE = 0.546$, $p = .277$).

Discussion

Our experimental results reveal regularisation behaviour in the production of complex systems of variation in morphology and word order. They also suggest that regularisation behaviour is of similar strength between these linguistic levels given input languages with comparable initial complexities. In Experiment 1 we found higher levels of regularisation in word order than in morphology, apparently due to the specific properties of the set of variants in the input languages. When both harmonic pre-nominal and post-nominal two-modifier variants were included, the coexistence of both variants in a single production system was rare. Although a preference for harmonic order and consistent head position may have been at play, the interference of L1 transfer cannot be categorically rejected. Indeed previous research suggests that L2 learners tend to access their L1 knowledge if it matches the novel input (Weber, Christiansen, Petersson, Indefrey, & Hagoort, 2016). In Experiment 2, we showed that eliminating opposite N-peripheral positions in the subset of two-modifier variants by replacing Num Adj N with N Num Adj eliminates the difference in regularisation between levels. Our results do not suggest general level-specific learning biases that could straightforwardly predict a typological asymmetry between the strength and speed of regularisation in morphology and word order hinted at in pidgin and creole studies (Good, 2015). Instead, they suggest that asymmetries in regularisation processes in language formation ought to be sought in asymmetries in the input complexity of traits across levels, also taking into account the overlap of features between contributing languages.

Conclusion

Our results suggest similar strengths of regularisation between linguistic levels given input languages with comparable initial complexities. Nevertheless, preferences for certain patterns within a linguistic level might in fact vary the strength of regularisation behaviour within a given level.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Culbertson, J., & Kirby, S. (2016). Simplicity and specificity in language: Domain-general biases have domain-specific effects. *Frontiers in Psychology*, *6*, 1964.

Culbertson, J., Smolensky, P., & Legendre, G. (2012). Learning biases predict a word order universal. *Cognition*, *122*(3), 306–329.

Fehér, O., Wonnacott, E., & Smith, K. (2016). Structural priming in artificial languages and the regularisation of unpredictable variation. *Journal of Memory and Language*, *91*, 158–180.

Ferdinand, V., Kirby, S., & Smith, K. (2017). The cognitive roots of regularization in language. *arXiv preprint arXiv:1703.03442*.

Good, J. (2015). Paradigmatic complexity in pidgins and creoles. *Word Structure*, *8*(2), 184–227.

Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language learning and development*, *1*(2), 151–195.

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

Kuznetsova, A., Bruun Brockhoff, P., & Haubo Bojesen Christensen, R. (2015). lmertest: Tests in linear mixed effects models [Computer software manual]. Retrieved from `http://CRAN.R-project.org/package=lmerTest`

R Core Team. (2015). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Senghas, A., & Coppola, M. (2001). Children creating language: How nicaraguan sign language acquired a spatial grammar. *Psychological science*, *12*(4), 323–328.

Siegel, J. (2004). Morphological simplicity in pidgins and creoles. *Journal of Pidgin and Creole languages*, *19*(1), 139–162.

Slabakova, R. (2013). What is easy and what is hard to acquire in a second language. *Contemporary approaches to second language acquisition*, *9*, 5.

St Clair, M. C., Monaghan, P., & Ramscar, M. (2009). Relationships between language structure and language learning: The suffixing preference and grammatical categorization. *Cognitive Science*, *33*(7), 1317–1329.

van Trijp, R. (2013). Linguistic assessment criteria for explaining language change: A case study on syncretism in german definite articles. *Language Dynamics and Change*, *3*(1), 105–132.

Weber, K., Christiansen, M. H., Petersson, K. M., Indefrey, P., & Hagoort, P. (2016). fmri syntactic and lexical repetition effects reveal the initial stages of learning a new language. *Journal of Neuroscience*, *36*(26), 6872–6880.

Wonnacott, E., & Newport, E. L. (2005). Novelty and regularization: The effect of novel instances on rule formation. In *Bucld 29: Proceedings of the 29th annual boston university conference on language development* (pp. 663–673).