Regularization of Word Order in the Verb Phrase differs from the Noun Phrase: Evidence from an online silent gesture perception paradigm

Monica Do (monicado@uchicago.edu)
Department of Linguistics, University of Chicago
1115 E. 58th Street, Chicago, IL 60637

Simon Kirby (simon.kirby@ed.ac.uk)
Centre for Language Evolution, University of Edinburgh
3 Charles Street, Edinburgh EH8 9AD, UK

Susan Goldin-Meadow (sgsg@uchicago.edu)
Department of Psychology, University of Chicago
5848 S. University Avenue, Chicago, IL 60637

Abstract

Prior work has shown a "natural" preference in the Verb Phrase for direct object Nouns to linearly precede the Verb. There is also evidence of a "natural" preference in the Noun Phrase to order Nouns before Adjectives. Given this, we asked how domain-general biases like regularization and language-specific biases like the preference for "natural" orders could jointly contribute to the emergence of these two common word orders cross-linguistically. Using a silent gesture paradigm (in which we presented iconic gestures without speech), we exposed different participants to competing Verb Phrase (NounVerb vs. VerbNoun) and Noun Phrase (NounAdj vs. AdjNoun) word orders at varying frequencies. In Noun Phrase contrast conditions, we found that regularization was greatest when the domain-general bias towards regularization and the linguistic bias to order Nouns before Adjectives were aligned. In Verb Phrase conditions, participants regularized to the same extent regardless of input: They opted for greater regularity, even at the expense of aligning with underlying word order biases. We discuss the implications of our work for understanding the effects of domain-general biases on language.

Keywords: Regularization, Frequency, Learning Biases, Gestures, Word Order, Adjective Order

Introduction

In order to talk about an event going on in the world, speakers have to make choices – they have to decide what to mention or not mention about an event as well as *how* to talk about the event. In particular, they have to determine the order in which to mention the different components – the Agent, the Patient, or the Action – of the event. Although there are a number of different ways in which they can do this, speakers in most of the world's languages tend to do so using one of two orders – the SOV or SVO order (Dryer, 2013b).

These two word orders are represented to similar degrees among the current languages of the world. However, there is evidence to suggest that there is an underlying cognitive bias towards the SOV order (e.g., Goldin-Meadow, So, Ozyurek, & Mylander, 2008; Langus & Nespor, 2010; Futrell, Hickey, Lee, Lim, Luchkina, & Gibson, 2015). Using the silent gesture paradigm, in which hearing participants are asked to communicate what happened in an event using only their hands, initial work by Goldin-Meadow et al. (2008) showed that, when gesturing about simple transitive events (e.g., a woman twisting a knob), speakers overwhelmingly mentioned the arguments in the Verb Phrase using the NounVerb, rather than the VerbNoun, order: Regardless of language background, people expressed the direct object, i.e., the Noun, before expressing the Verb. Moreover, in line with these results, the world's newest languages, including Al-Sayyid Bedouin Sign Language (Sandler, Meir, Padden, & Aronoff, 2005), and the homesigning systems of deaf individuals (Goldin-Meadow & Mylander, 1998) have also tended towards the SOV order. Later work using the same silent gesture paradigm has extended those findings, showing that the "natural" preference for (object) NounVerb word order may be modulated by additional factors – such as the animacy (Meir et al., 2017), reversibility (Gibson et al., 2013; Hall, Mayberry, & Ferreira, 2013), and whether the event is intensional or extensional (Schouwstra & de Swart, 2014). Nevertheless, there is general agreement that NounVerb order is expected for prototypical, animate-inanimate events.

In the Noun Phrase, a similar picture has emerged: Some initial evidence has pointed toward an underlying cognitive bias in the ordering of Nouns and their modifying Adjectives. Cross-linguistically, for instance, roughly 64% of the world's spoken languages order Nouns before Adjectives compared to the roughly 27% with the reverse order (Dryer, 2013a). Recent work has also shown the NounAdj order to predominate in signed languages (Coons, 2022). In addition, there has been some initial evidence for a NounAdj bias from silent gesture studies by Culbertson at al. (2020), who reported that the two most frequent word orders among English silent gesturers to be the Dem-Num-Noun-Adj and the Dem-Noun-Adj-Num orders. Although the bias to order Nouns before Adjs is certainly deserving of further investigation, these results do nonetheless provide strong initial support for a similar "natural" preference in the ordering of the Noun Phrase. Taken together, then, studies using the silent gesture paradigm have shown that cognitive biases in the non-verbal representation of events as well as objects can play an important role in motivating the word order of languages, more generally.

Building on these initial findings, more recent work has argued that broad typological tendencies in language are unlikely to be the product of domain-specific biases like the

preference for "natural" orders alone. Instead, these studies have argued that broader cross-linguistic tendencies are attributable to the way that domain-general cognitive biases – in particular, the well-known bias towards regularization (Hudson Kam & Chang, 2009; Hudson Kam & Newport, 2005, 2009; Singleton & Newport, 2004; Smith & Wannacott, 2010) – operate on biases specific to the domain of language (Culbertson & Kirby, 2016; Ferdinand et al., 2019; Motamedi et al., 2021a; Motamedi, Wolters, Schouwstra et al., 2021; Saldana, Smith, Kirby, & Culbertson, 2021).

The Current Study

The aim of the present work is, therefore, to further understand how domain-general biases such as the bias towards regularization interact with two different ordering biases in language: the ordering of elements inside the Verb Phrase and the Noun Phrase. To do this, we use the silent gesture regularization paradigm initially introduced by Motamedi et al. (2021a). In this paradigm, participants are initially shown an event and then trained on two competing word orders (e.g., NounVerb versus VerbNoun or NounAdj versus AdjNoun) that describe the event using only gesture. In a departure from traditional silent gesture studies in which participants spontaneously produce their own gestures, we – following Motamedi et al. (2021a) – ask participants to complete a forced choice task where they select between the two training orders and see which of the two orders (if any) is over-produced, i.e., regularized, at test.

We expect participants to regularize to a greater extent when the bias to regularize coincided with their "natural" ordering preferences (Goldin-Meadow et al., 2008). When contrasting the order of the Noun versus the Verb, for instance, we thus expect participants to regularize to a greater extent when majority of their input is NounVerb, compared to when it is VerbNoun. Likewise, when contrasting the order of Nouns versus Adjectives, we expect participants to regularize more in conditions where the majority of their input is NounAdj compared to cases where it is AdjNoun. We also compare rates of regularization in the Verb Phrase versus Noun Phrase contexts, though predictions here are less clear.

Methods

Participants Data from 320 native English-speakers recruited from Prolific were submitted to analysis. They were all at least 18 years of age and reported little to no prior experience with sign language.

Materials We created four different event clips. Each clip was 3-4 seconds in length and depicted a person performing a simple action (e.g., a person waving a zigzag patterned spoon back and forth). Three of these event clips were shown to participants during an initial training phase while a fourth, held-out event was only shown during the test phase.

Each event clip was accompanied by four distinct gesture vignettes. These gesture vignettes always consisted of three gestures: a Noun-denoting gesture (i.e., spoon), an Adjective-denoting gesture (i.e., zigzag), and a Verb-denoting gesture

(i.e., motion back and forth). The gestures and timing of these vignettes was identical except for the order in which the gestures occurred. Specifically, gestures were performed in one of four orders: (i) the Adj-Noun-Verb (ANV), (ii) Noun-Adj-Verb (NAV), (iii) Verb-Adj-Noun (VAN), or (iv) Verb-Noun-Adj (VNA).

Procedure The structure of the experiment was the same for both the NounVerb and NounAdj contrast conditions. During an initial training phase, participants were exposed to an event clip followed by a gesture vignette. Each event clip was seen four times, each time paired with one of two different gesture vignettes. Critically, though, the frequency with which they saw each of these different vignettes differed by condition. In the Majority NounFirst condition, participants saw the NounFirst order in 75% of the training trials for any given event and the alternative NounLast order in 25% of the training trials for that event. In the Majority NounLast condition, these frequencies were reversed. In total, participants were shown 12 event-gesture pairs.

During the test phase, participants were shown an event and asked to decide which of two orders best described the event. Participants completed 16 total test trials; twelve of these included the exact same event-gesture pairs previously seen during training while an additional four trials required them to select between one of two gesture vignettes depicting a completely new, previously unseen event clip (Kirby et al., 2008; Motamedi et al., 2021a; Motamedi et al., 2021b).

Table 1: List of Conditions

Contrast	Majority Order	MajOrder (75%) - MinOrder (25%)
NounVerb	NounFirst	NAV-VNA
NounVerb	NounFirst	ANV-VAN
NounVerb	NounLast	VNA-NAV
NounVerb	NounLast	VAN-ANV
AdjNoun	NounLast	ANV-NAV
AdjNoun	NounFirst	NAV-ANV
AdjNoun	NounLast	VAN-VNA
AdjNoun	NounFirst	VNA-VAN

We thus manipulated two factors, ContrastType (AdjNoun vs NounVerb) and Majority Order (Majority NounFirst vs Majority NounLast), producing a total of four distinct between-subject testing conditions. But, because our gesture sequences consisted of three elements, we created an additional four conditions, which varied only the order of the non-target element in the sequence. We tested, for instance, NounAdj ordering preferences both when the non-target verb element followed the entire Noun Phrase (ANV-NAV) and when it preceded the Noun Phrase (VAN-VNA). This produced a total of eight different between-subject conditions to which participants were randomly assigned (Table 1).

Results

The degree to which participants regularized their initial input was quantified via change in Shannon's entropy (Shannon, 1948), which we calculated by subtracting the input entropy from the output entropy for each. The input entropy, the entropy score for the original training stimuli, was always .81 bits because vignettes were always shown in a 3:1 ratio. The output entropy for each participant was determined based on the ratio at which the two vignette variants were selected at test. Thus, negative Entropy Change scores indicate less variation at test than during training with smaller (i.e., more negative) scores corresponding to greater degrees of regularization at test.

NounVerb Order Figure 1 (Left) shows the mean Entropy Change score for each of the NounVerb Contrast conditions. For each condition, we first ran an intercept-only linear model to see whether there was any evidence of regularization at all. In line with what is shown in Figure 1, Entropy Change scores were significantly less than zero in all conditions meaning that there was more regularity (p's < .001) in participants' responses than in the initial training stimuli.

When we compared rates of regularization across conditions, though, we failed to find any significant main or interaction effects relating to either Majority Order (p's > .2) or to the location of the non-target adjective (p's > .2). This suggested participants in the NounVerb Contrast conditions regularized to the same extent, regardless of input.

To take a closer look at which orders participants were regularizing towards, we also analyzed the proportion of Majority Order selections made by participants at test (Figure 2, Left). These analyses yielded no significant main or

interaction effects involving either Majority Order (i.e., NounFirst versus NounLast) or adjective location (p's > .09), meaning that participants selected the Majority Order vignette to the same extent regardless of whether the majority of their input as NounFirst or NounLast and regardless of the location of the non-target adjective. The proportion of Majority and Minority Order selections for each condition in the NounVerb Contrast is given in Table 2.

Table 2: Proportion of Selections by Condition in the NounVerb Contrast

ConditionID	Prop. Maj. Order Selections	Prop. Min. Order Selections
NAV-VNA	.72	.28
ANV-VAN	.59	.41
VNA-NAV	.66	.34
VAN-ANV	.71	.29

NounAdj Order Figure 1 (Right) shows Entropy Change scores in each of the NounAdj Contrast conditions. We again ran an intercept-only linear model to determine whether there were significant decreases in Entropy (i.e. variation) within each condition. We found significant Entropy reductions in only the conditions where the Majority Order was NounFirst (p's < .001); Majority NounLast conditions did not show any significant differences from zero (p's > .1), meaning that there was no evidence of regularization in these conditions.

Comparisons across conditions showed a similar pattern of results. There was a main effect of Majority Order ($\beta = 0.17$, SE = 0.06, |t| = 3.00, p < .01), such that participants
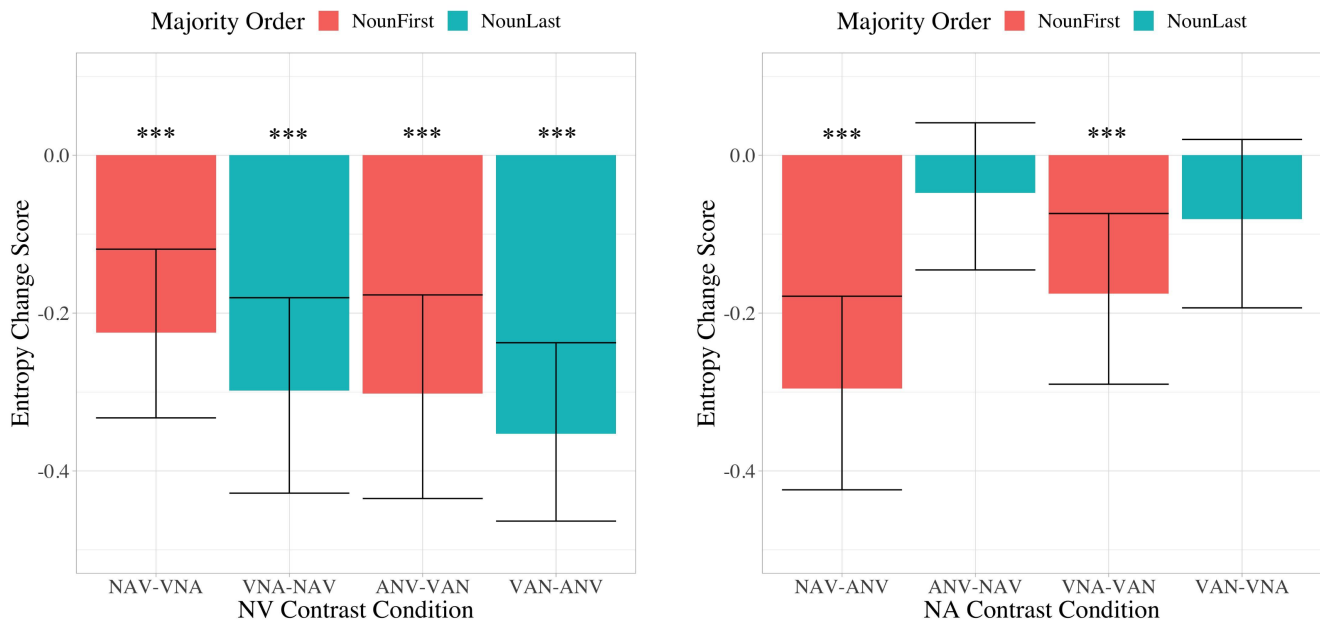

Figure 1 Mean Entropy Change scores in each condition of the NV Contrast (A) and NA Contrast (B). Majority Orders are listed first. Negative scores point to evidence of regularization. Error bars indicated bootstrapped 95% CIs.

regularized to a greater extent when the majority of the input that they received was NounFirst, compared to when the majority of their input was NounLast. This is true regardless of the position of the non-target Verb element: we failed to detect any main or significant interactions involving the location of the Verb (p's > .4). In other words, unlike in the NounVerb Contrast, there were significant differences in the degree to which participants regularized across conditions: In the NounAdj Contrast, participants regularized significantly more when the majority of their input was in the NounFirst order compared to when it was the NounLast order.

As before, we analyzed the proportion of Majority Order selections made by participants (Figure 2, Right). The proportion of Majority and Minority selections in each condition of the NounAdj Contrast is given in Table 3. Comparison across conditions revealed a significant main effect of Majority Order (β = -1.16, SE = 0.46, $|z|$ = 2.53, p < .05) such that participants were more likely to select Majority Order vignette when that vignette was in the NounFirst configuration. Effects involving Verb Location did not reach significance, suggesting that the Majority Order effect did not depend on the location of the non-target Verb element.

NounVerb vs NounAdj Comparison Finally, we compared the Entropy Change Scores between NounVerb and NounAdj Contrast conditions. Interestingly, we found a main effect of Contrast Type (β = 0.14, SE = 0.04, $|t|$ = 3.45, p < .001), such that participants regularized *more* in the NounVerb Contrast than in the NounAdj Contrast. This effect was modulated by a significant Contrast x Majority Order interaction (β = 0.23, SE = 0.08, $|t|$ = 2.78, p < .01) indicating that the difference between Contrast Types was largely driven by lower rates of regularization in the NounLast Majority Order conditions for the NounAdj Contrast.

Table 3: Table 2: Proportion of Selections by Condition in the NounAdj Contrast

ConditionID	Prop. Maj. Order Selections	Prop. Min. Order Selections
NAV-ANV	.78	.22
ANV-NAV	.57	.43
VNA-VAN	.73	.28
VAN-VNA	.63	.37

Discussion

The aim of this work was to further understand how domain general biases, such as the bias towards regularization, interact with word order biases in language to drive the development of NounVerb and NounAdj word orders cross-linguistically. To that end, we used a silent gesture perception study modeled after recent work by Motamedi et al. (2021a) and examine the degree to which participants would regularize the order of the elements in two different syntactic contexts – the Verb Phrase and the Noun Phrase.

In both cases, we predicted that speakers would regularize towards the dominant input order to a greater extent when that order coincided with the "natural" preference to order Nouns before Verbs (e.g., Goldin-Meadow et al., 2008) and Nouns before Adjectives (Culbertson et al., 2020).

This prediction was born out in the NounAdj condition. Here, Entropy Change scores showed that English-speaking participants did, indeed, regularize their input, but *only* when the majority of their input during training was in the NounFirst configuration (e.g., in NAV or VNA orders) – the order that appears most frequently across the world's languages (Dryer, 2013). We did not, by contrast, find any
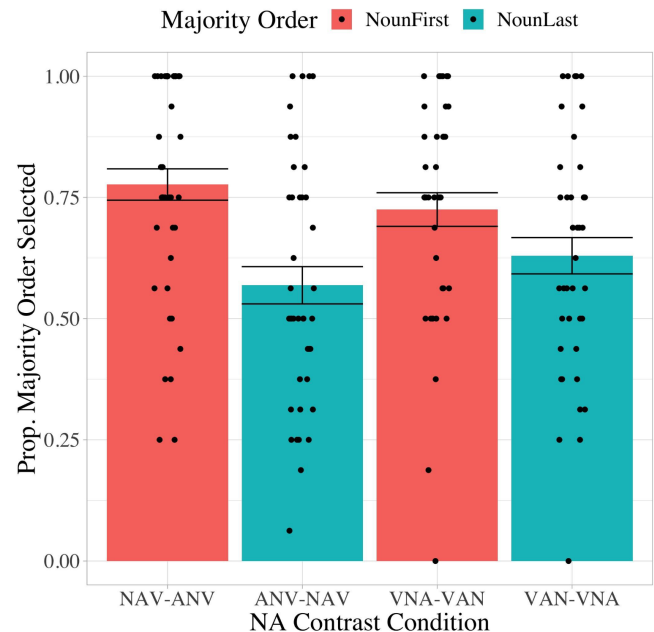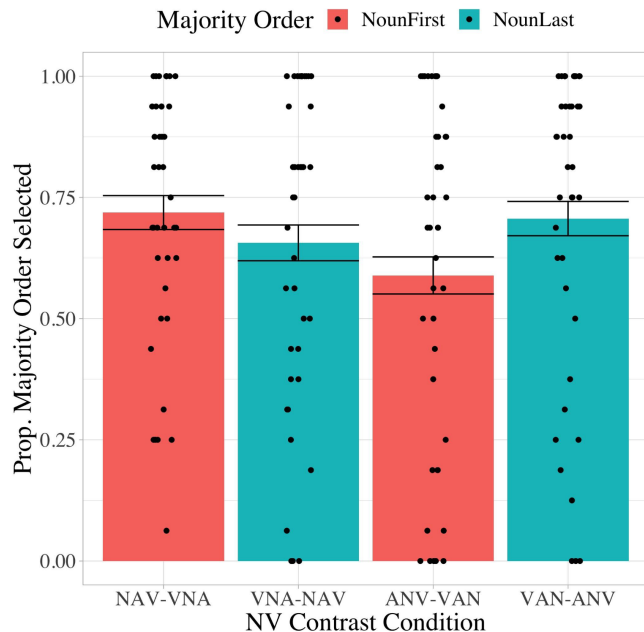

Figure 2 Mean proportion of Majority Order selections for each condition of the NounVerb (Left) and NounAdj (Right) Contrasts. Each individual point represents a single participant. Majority Orders are listed first. Error bars indicate 95% CI.

evidence of regularization when the majority of their input was in the NounLast configuration (e.g., in ANV or VAN orders), even though this was the more English-like word order. Rather, analyses over the proportion of Majority versus Minority orders selected showed that the reason why we failed to find evidence of regularization in these latter conditions was because participants in these conditions were overproducing the typologically dominant NounFirst order, not the Majority NounLast order. That is, they selected the typologically dominant, "natural" NounFirst order in over 55% of trials, even though these orders comprised only 25% of their input.

Our results also add to the growing body of work (Culbertson & Kirby, 2016; Ferdinand et al., 2019; Motamedi et al, 2021a) geared towards understanding the potentially different ways in which domain-general biases are operationalized in linguistic contexts. In particular, results from the NounAdj Condition suggest that one important contributor to the prevalence of NounAdj word orders, cross-linguistically, may be the presence of a domain-general biases like the bias towards regularization acting in confluence with the more specific word order biases that might exist when people are engaged in a linguistic task.

Likewise, the lack of regularization in the Majority NounLast conditions suggests that the presence of typologically less frequent orders like the AdjNoun order may be the product of domain-general biases working in competition against language-specific biases. Given that AdjNoun orders do nevertheless appear in roughly 27% of languages, though, an interesting avenue for future work is the contexts under which participants in the Majority NounLast conditions would eventually regularize towards the AdjNoun order.

In the NounVerb Contrast conditions, however, we found a different pattern of results. Unlike in the NounAdj Contrast conditions, where participants regularized their input in *only* the "natural" Majority NounFirst conditions, participants in the NounVerb Contrast conditions regularized towards the majority order to the same extent in both the Majority NounFirst and Majority NounLast conditions. That is, participants extended the majority order both when that order was in the "natural" NounFirst configuration and when it was in the "unnatural" NounLast configuration.

The fact that participants actually regularized their input to a *greater* extent in the NounVerb than in the NounAdj Condition points to an interesting possibility – specifically that the desire to achieve regularity in NounVerb word orders may overwhelm the more domain-specific bias towards any particular word order. If this is the case, then, an important avenue for future research will be determining precisely *why* participants were more willing to violate their "natural" ordering preferences in NounVerb compared to NounAdj conditions.

Here, one possibility may be that the strength of the bias for the NounAdj order in the Noun Phrase is simply stronger than the strength of the bias for NounVerb in the Verb Phrase. Some evidence for this comes from the fact that the overwhelming preference for the NounVerb order found in silent gestures production has not translated to the prominence of the SOV order cross-linguistically. Moreover, historically, there have been many examples of languages evolving from the SOV to what some have argued to be the more "stable" SVO order (e.g., Newmeyer, 2000; Bauer, 1995; Kiparsky, 1996; Leinonen, 1980; Fisher, 1975; among others). When compared to the clear-cut preference for the NounAdj in both silent gesture studies and among the world's languages, the parity of the SOV versus SVO orders cross-linguistically may suggest that the cognitive preference to order Nouns before Verbs may be more vulnerable than the preference to order Nouns before Adjs.

An altogether different possibility for the higher rates of regularization in the NounVerb versus NounAdj contrasts may be related to the status of Adjectives versus direct object Nouns in the conceptual representation of an object/event – a distinction that is reflected in the Adjunct versus Argument status of Adjectives versus direct objects syntactically. In particular, when describing transitive events like the ones used in our study, the object Noun is an inextricable argument of the Verb: one cannot simply "toss", one must toss something. Adjectives, by contrast, are not vital to the interpretation of the Noun. In this case, one reason why participants may have been more willing to violate their more natural preferences for the sake of regularity may be related to differences in the conceptual relationship between a Verb and its direct object versus a Noun and its Adjective, rather than in the strength of the bias, itself.

Complicating both of these accounts, though, is the open issue of why participants in the NounVerb Contrast of our study demonstrated an unexpected willingness to also regularize the "unnatural" NounLast (i.e., VerbNoun) word orders – especially given that participants in Motamedi et al.'s study (2021a; extensional conditions), which used the same paradigm that we did here, did no such thing. Here, one possibility may be the way in which events were depicted in our study versus in Motamedi et al. (2021a). Specifically, whereas participants in our study were shown videos depicting dynamic motion, participants in Motamedi et al., (2021a) were shown static images representing each of the events. In addition, to reduce the complexity of the task for our participants, we chose to "background" the entity performing the action by making sure that their face was out of frame in the event clips (and by omitting the Subject from gesture sequences).[1] The same was not true for the stimuli used in case for Motamedi et al. (2021a); in their stimuli, the entity performing the action was always apparent in the images used to depict the event. It is possible, then, that

---

[1] We thank an anonymous reviewer for pointing out separate ways in which "backgrounding" the Agent in the event may change the way in which the event is interpreted and then subsequently described. While we cannot fully rule out the possibility that

"backgrounding" the Agent led viewers to construe the spoon item as the Subject of the event (e.g., as 'a spoon moving back and forth'), initial data gathered from a written production version of this study does not seem to support this possibility.

participants in our study were particularly willing to regularize the "unnatural" NounLast order because the NounLast – i.e., VerbNoun – order allowed them to mention the most salient aspect of the event, the motion, first. While additional work is required to investigate the role that this type of salience might play on the bias towards ordering the direct object Noun before the Verb, this account does appear to be broadly consistent with prior work looking at the ways in which the natural preference for the NounFirst order can be modulated (Meir et al., 2017; Gibson et al., 2013; Hall, Mayberry, & Ferreira, 2013; Schouwstra & de Swart, 2014).

Conclusion

Extending work by Culbertson et al. (2020), the studies presented here introduced a direct comparison between NounFirst versus NounLast word orders and provided additional evidence that the prevalence of the NounAdj order cross-linguistically may stem from the bias towards regularization acting in conjunction with the underlying cognitive bias towards the NounAdj order. An important avenue for future work may be investigating the different speaker-internal (Culbertson et al., 2020; Hall et al., 2013) and speaker external factors (Gibson et al., 2013) ultimately underlying the bias for post-nominal adjectives.

At the same time, our results appear to pose an interesting exception to prior work (e.g., Saldana et al., 2021) pointing towards uniformity in the strength of the regularization bias across levels of linguistic structure (e.g., morphological versus syntactic structure). In particular, differences in the extent to which participants were willing to regularize competing word orders in our NounVerb versus NounAdj Contrasts suggest that the domain-general bias towards regularization may also depend on the type of construction. We believe that an interesting avenue of future work will be investigating why this may be the case.

References

Bauer, B. L. M. (1995). *The emergence and development of SVO patterning in Latin and French*. Oxford: Oxford University Press.

Coons, C. (2022). Nominal Word Order Typology in Signed Languages. *Frontiers in Communication*, 6, 802596. https://doi.org/10.3389/fcomm.2021.802596

Culbertson, J., Schouwstra, M., & Kirby, S. (2020). From the world to word order: Deriving biases in noun phrase order from statistical properties of the world. *Language*, *96*(3), 696–717. https://doi.org/10.1353/lan.2020.0045

Dryer, M. S. (2013a). Order of Adjective and Noun. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology. https://wals.info/chapter/87

Dryer, M. S. (2013b). Order of Object and Verb. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology. https://wals.info/chapter/83

Ferdinand, V., Kirby, S., & Smith, K. (2019). The cognitive roots of regularization in language. *Cognition*, *184*, 53–68. https://doi.org/10.1016/j.cognition.2018.12.002

Fischer, S. (1975). Influences on word order change in American Sign Language. In C. N. Li (Ed.) *Word Order and Word Order Change*. Austin: University of Texas Press.

Futrell, R., Hickey, T., Lee, A., Lim, E., Luchkina, E., & Gibson, E. (2015). Cross-linguistic gestures reflect typological universals: A subject-initial, verb-final bias in speakers of diverse languages. *Cognition*, *136*, 215–221. https://doi.org/10.1016/j.cognition.2014.11.022

Gibson, E., Piantadosi, S. T., Brink, K., Bergen, L., Lim, E., & Saxe, R. (2013). A Noisy-Channel Account of Crosslinguistic Word-Order Variation. *Psychological Science*, *24*(7), 1079–1088. https://doi.org/10.1177/0956797612463705

Goldin-Meadow, S., & Mylander, C. (1998). Spontaneous sign systems created by deaf children in two cultures. *Nature*, 391(6664), 279–281. https://doi.org/10.1038/34646

Goldin-Meadow, S., So, W. C., Ozyurek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences*, *105*(27), 9163–9168. https://doi.org/10.1073/pnas.0710060105

Hall, M. L., Mayberry, R. I., & Ferreira, V. S. (2013). Cognitive constraints on constituent order: Evidence from elicited pantomime. *Cognition*, *129*(1), 1–17. https://doi.org/10.1016/j.cognition.2013.05.004

Hudson Kam, C. L., & Chang, A. (2009). Investigating the cause of language regularization in adults: Memory constraints or learning effects? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(3), 815–821. https://doi.org/10.1037/a0015097

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. https://doi.org/10.1080/15475441.2005.9684215

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. https://doi.org/10.1016/j.cogpsych.2009.01.001

Kiparsky, P. (1996). The shift to head-initial VP in Germanic. In H. Thrainsson, J. Peter, & S. Epstein (Eds.), *Comparative Germanic syntax*. Kluwer.

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, *105*(31), 10681–10686. https://doi.org/10.1073/pnas.0707835105

Langus, A., & Nespor, M. (2010). Cognitive systems struggling for word order. *Cognitive Psychology*, *60*(4), 291–318. https://doi.org/10.1016/j.cogpsych.2010.01.004

Leinonen, M. (1980). A closer look at natural serialization. *Nordic Journal of Linguistics*, 3, 147–159. https://doi.org/10.1017/S033258650000055X

Meir, I., Aronoff, M., Börstell, C., Hwang, S.-O., Ilkbasaran, D., Kastner, I., Lepic, R., Lifshitz Ben-Basat, A., Padden, C., & Sandler, W. (2017). The effect of being human and the basis of grammatical word order: Insights from novel communication systems and young sign languages. *Cognition*, 158, 189–207. https://doi.org/10.1016/j.cognition.2016.10.011

Motamedi, Y., Wolters, L., Naegeli, D., Schouwstra, M., & Kirby, S. (2021a). Regularisation, Systematicity and Naturalness in a Silent Gesture Learning Task. In Proceedings of the Annual Meeting of the Cognitive Science Society (Vol. 43, No. 43).

Motamedi, Y., Wolters, L., Schouwstra, M., & Kirby, S. (2021b). *The effects of iconicity and conventionalisation on word order preferences* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/u5amg

Newmeyer, F. J. (2000). On the Reconstruction of 'Proto-World' Word Order. In C. Knight, J. R. Hurford, & M. Studdert-Kennedy (Eds.), *The evolutionary emergence of language: Social function and the origins of linguistic form*. Cambridge University Press: Cambridge.

Saldana, C., Smith, K., Kirby, S., & Culbertson, J. (2021). Is Regularization Uniform across Linguistic Levels? Comparing Learning and Production of Unconditioned Probabilistic Variation in Morphology and Word Order. *Language Learning and Development*, 17(2), 158–188. https://doi.org/10.1080/15475441.2021.1876697

Sandler, W., Meir, I., Padden, C., & Aronoff, M. (2005). The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences*, 102(7), 2661–2665. https://doi.org/10.1073/pnas.0405448102

Schouwstra, M., & de Swart, H. (2014). The semantic origins of word order. *Cognition*, 131(3), 431–436. https://doi.org/10.1016/j.cognition.2014.03.004

Senghas, A., & Coppola, M. (2001). Children Creating Language: How Nicaraguan Sign Language Acquired a Spatial Grammar. *Psychological Science*, 12(4), 323–328. https://doi.org/10.1111/1467-9280.00359

Shannon, C. E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Singleton, J. L., & Newport, E. L. (2004). When learners surpass their models: The acquisition of American Sign Language from inconsistent input. *Cognitive Psychology*, 49(4), 370–407. https://doi.org/10.1016/j.cogpsych.2004.05.001

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