

# Do cross-linguistic patterns of morpheme order reflect a cognitive bias?

**Carmen Saldana (carmen.saldana@ed.ac.uk)**

Centre for Language Evolution, The University of Edinburgh,  
3 Charles Street, Edinburgh, EH8 9AD, UK

**Yohei Oseki (oseki@aoni.waseda.jp)**

Faculty of Science and Engineering, Waseda University,  
3-4-1 Okubo, Shinjuku, Tokyo 169-8555, Japan

**Jennifer Culbertson (jennifer.culbertson@ed.ac.uk)**

Centre for Language Evolution, The University of Edinburgh,  
3 Charles Street, Edinburgh, EH8 9AD, UK

## Abstract

A foundational goal of linguistics is to investigate whether shared features of the human cognitive system can explain how linguistic patterns are distributed across languages. In this study we report a series of artificial language learning experiments to test a hypothesised link between cognition and a persistent regularity of morpheme order: number morphemes (e.g., plural markers) tend to be ordered closer to noun stems than case morphemes (e.g., accusative markers) (Greenberg, 1963). We argue that this typological tendency may be driven by a bias favouring orders that reflect scopal relationships in morphosyntactic composition (Bybee, 1985; Rice, 2000; Culbertson & Adger, 2014). We taught participants an artificial language with noun stems, and case and number morphemes. Crucially, the input language indicated only that each morpheme preceded or followed the noun stem. Examples in which two (overt) morphemes co-occurred were held out—i.e., no instances of plural accusatives. At test, participants were asked to produce utterances, including the held-out examples. As predicted, learners consistently produced number closer to the noun stem than case. We replicate this effect with free and bound morphemes, pre- or post-nominal placement, and with English and Japanese speakers. However, we also find that this tendency can be reversed when the form of the case marker is conditioned on the noun, suggesting an influence of dependency length. Our results provide evidence that universal features of cognition may play a causal role in shaping the relative order of morphemes.

**Keywords:** linguistic universals; artificial language learning; morpheme order; case; number

## Introduction

Human languages are incredibly diverse in the way they combine meaningful units, i.e., morphemes; nevertheless, certain regularities are apparent. For example, some patterns of morpheme order occur more frequently across the languages of the world, while others are rare or even unattested. The typological regularity in morpheme order we target here concerns number and case morphology, specifically, languages in which there is a boundary between these morphemes. For example, in agglutinating languages such as Hungarian or Turkish, there is distinct set of number morphemes (marking plurality) and case morphemes (marking grammatical roles). In such languages, when overt morphemes of both number and case are present on a stem, and both follow or both precede the noun stem, the expression of number is almost always realised closer to the noun stem than the expression of case (Universal 39; Greenberg, 1963). There

are a number of candidate explanations for this phenomenon, which intersect with high-level hypotheses about how morpheme (and word) order is determined in language more generally. For example, it has been proposed that semantic or compositional relationships among morphemes, sometimes called *scope*, determine linear order (Bybee, 1985; Wunderlich, 1993; Rice, 2000; Culbertson & Adger, 2014).<sup>1</sup> On one formulation, morphemes which more directly affect or modify the semantic content of the stem have narrower scope (Bybee, 1985; Rice, 2000). Wider-scope morphemes modify the larger semantic constituent which includes any lower scoping morphemes. Perhaps the best-known example of this is the order of derivational and inflectional morphemes (e.g., ‘neighbor-hood-s’). On this account, derivational morphemes are ordered closer to the stem because they change its lexical meaning. Inflectional morphemes scope higher, modifying grammatical properties of the stem plus any derivational morphemes. Similarly, it has been claimed that the linear order of nominal modifiers (e.g., adjectives, numerals, demonstratives) reflects semantic scope relations (Culbertson & Adger, 2014; Bouchard, 2002). In the case of Universal 39, the idea would be that case scopes higher than number because number directly modifies the entity referred to by the noun, while the case morpheme signals an external relationship between the entity and some event. Following Culbertson and Adger (2014), we call orders which reflect scope relations *scope-isomorphic*.

A second possible explanation appeals to frequency and its effects on processing. For example, Ryan (2010) shows that in some cases morpheme order reflects the frequency of stem+morpheme bigrams (see also Baayen, 1993; Rice, 2011). Along similar lines, Hay (2001) argues that when a stem is more frequent alone than with a particular affix, then that affix is easier to parse (decompose) from the stem. This in turn determines linear order: more parsable affixes appear farther from the stem than less parsable ones (see also Hay & Plag, 2004; Plag & Baayen, 2009; Manova & Aronoff, 2010). How might this explain Universal 39? It could be that

<sup>1</sup>Related theories argue that universal morphosyntactic hierarchies, potentially reflecting semantics, determine order (Baker, 1985; Grimshaw, 1986; Cinque, 2005).

number tends to be expressed more often than case, or that case morphemes tend to be more parsable than number morphemes. On this account, there is nothing about the semantics of these morphemes that determines their relative order. Indeed, a third possibility is that their relative order reflects patterns of diachronic change: it could be that languages tend to grammaticalise number before case (Givón, 1979).

To date, there is no direct behavioral evidence adjudicating among these potential explanations for Universal 39. In fact, there is no independent evidence beyond the typology to show that placing number closer to the noun stem than case is in fact preferred over the reverse. In a series of three artificial language learning experiments, we test the link between this typological generalisation and a bias towards linear orders that mirror scopal relationships (henceforth scope-isomorphic orders). To summarise, we find support for this hypothesis across two language populations (English, and Japanese) independent of morpheme position (before or after the noun stem), degree of boundedness, and frequency. All things equal, learners therefore prefer scope-isomorphic orders. However, we also find that conditional allomorphy between the stem and the case marker can reverse participants' preferences. We interpret this as a competing bias for local dependencies. This result adds to the growing body of work using these experimental methods to investigate how learning and use shape morphology and word order (Hupp, Sloutsky, & Culicover, 2009; Fedzechkina, Jaeger, & Newport, 2012; Culbertson & Adger, 2014; Culbertson, Smolensky, & Legendre, 2012; Tabullo et al., 2012; Futrell, Mahowald, & Gibson, 2015; Fedzechkina, Chu, & Jaeger, 2018).

## Experiment 1

### Methods

The artificial language learning experiments described here use an extrapolation paradigm (called 'Poverty-of-the-stimulus' paradigm elsewhere, Wilson, 2003; Culbertson & Adger, 2014). This means learners are trained on input that is designed to be ambiguous between (at least) two patterns of interest: here, two potential ways of ordering case and number morphemes. Learners are exposed to a miniature artificial language with nouns, and case (accusative) and number (plural) morphemes. Crucially, their input indicates whether these morphemes generally precede or follow the noun, but does not include any examples in which the two morphemes co-occur within the same noun phrase. At test, they are asked to produce utterances, including these held out examples. The order they infer will indicate whether they have a preference for placing number closest to the noun (e.g., Noun-Number-Case rather than Noun-Case-Number). All experiment materials and data discussed here are available at [osf.io/9fa3v/](https://osf.io/9fa3v/), and the preregistered design and analysis plan for Experiment 1 is accessible at [osf.io/8xuc9](https://osf.io/8xuc9).

**Participants** Forty-one native English speakers were recruited from the University of Edinburgh's Careers Services database. Participants were paid £6 for a 35-min-long exper-

imental session. Participants (N=1) whose vocabulary accuracy was lower than 60% were excluded; testing trials with incomplete sentences were also excluded.

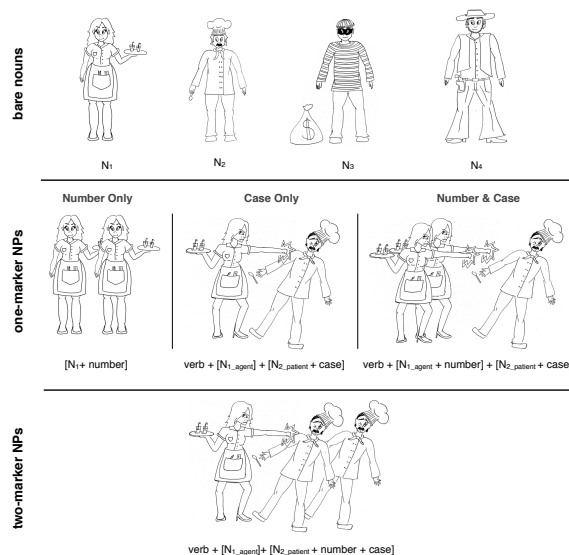


Figure 1: Example visual stimuli and corresponding descriptions. Top to bottom: the four characters in the miniature language in isolation; example events with one marker (either number or case); example event requiring two markers (number and case, testing only).

**Input language** The lexicon includes three semi-nonce verbs, four nonce nouns, and two nonce markers (one number marker indicating plural; one case marker indicating accusative). All words have initial stress. The three semi-nonce verbs are taken from the English-based creole Tok Pisin: 'kikim' ([k'ikim]), 'poinim' ([p'ɔim]) and 'straikim' ([strakim]), which refer to 'kicking', 'pointing' and 'punching' respectively. The (disyllabic) nouns are 'negid' ([neʒid]), 'nork' ([nɔrk]), 'tumbat' ([t'ʰʌmbət]), 'vaem' ([væm]) (based on Fedzechkina et al., 2012), naming four characters: a burglar, a chef, a cowboy, and a waitress. The noun-character mappings are random for each participant. The two markers were randomly mapped to number and case from the set: 'gu' ([gu:],), 'sa' ([sɑ:],) and 'ti' ([tʰi:]). Word order in sentences was Verb-Agent-Patient. Half of participants were trained on a language with post-nominal morphemes (case and number morphemes appeared after the noun stem), half with pre-nominal morphemes (case and number morphemes appeared before the noun stem).<sup>2</sup>

Participants are trained on three different NP types: a bare noun, a noun with overt number morphology, and a noun with

<sup>2</sup>We use the terms *pre-* and *post-nominal* instead of *prefixal* and *suffixal* morphology to account for both bound and unbound orthographic representations of case and number morphology.

overt case morphology. Note that singular, and agent case (nominative) are unmarked. During training, participants get descriptions of characters in isolation (singular or plural), or events with a singular patient; plural patients (requiring both number and case morphology) are held-out until testing. See Figure 1 for examples. Crucially, number and case markers appear with the exact same frequency (i.e., absolute, and relative to each given noun) both during training and testing phases, controlling for any potential frequency effects.

The input language is presented both orthographically and auditorily during training. Auditory stimuli were recorded in a sound-attenuated room by a 26yo male speaker of American English. Noun phrases were recorded without a pause between nouns and markers but each marker is orthographically presented surrounded by spaces and thus not bound to the noun.

**Experimental procedure** The experiment was conducted in a quiet room, with all instructions provided in English, and an English-speaking experimenter. Participants were told that they would be learning part of a foreign language. The session proceeded as follows.

**Phase 1, noun training and testing.** Participants are first trained on the four nouns in isolation (Figure 1, top row) during a block of 24 trials (6 per noun). In each trial, a single character appears, and its description (a bare noun) is displayed (orthographically and auditorily). Participants are instructed to repeat each description aloud. Participants are then tested on the noun vocabulary using a noun-selection task and an oral production task (12 trial per block, 3 per noun). In noun-selection trials, a character appears, and participants must select the correct noun from 2 choices. The foil noun is randomly selected at each trial. Feedback is provided (an (in)correct-answer sound effect along with the image and correct noun; if incorrect, the audio of the noun is also played). In oral production trials, a character appears, and participants must say the corresponding noun aloud. Feedback is provided (the correct noun is displayed visually and auditorily after participants submit their answer).

**Phase 2, one-marker NP training.** Participants are next trained on noun phrases with a single marker, either number or case. There are three trial types (Figure 1, middle row): (1) a group of the same characters (2, 3, or 4) in isolation (Number only), (2) an event with (different) singular agent and patient (Case only), or (3) an event with a plural agent, and a singular patient (Number & Case, where crucially each marker belongs to a different noun phrase). On each training trial, participants see an image, and its description is presented (orthographically and auditorily). There are 62 trials total (randomised): 8 bare noun, 18 Number Only (six per character), 18 Case Only (randomly chosen from the 36 possible), and 18 Number & Case images (again randomly chosen).

**Phase 3, one-marker NP comprehension test.** Participants are then tested on their comprehension of one-marker NPs in an image-selection task. On each trial, they get a description

and must select the corresponding image out of an array of two. Feedback is provided (an (in)correct-answer sound effect along with the image and correct orthographic description; if incorrect, the audio description is also played). The foil image is selected according to the trial type. For bare noun and Number Only trials, the foil image is the same character with wrong numerosity (e.g., singular instead of plural). For Case Only and Number & Case trials, the foil is the same event type with agent and patient reversed. There are 34 trials total (randomised): 4 bare noun, 10 each of the three one-marker NP trial types.

**Phase 4, one-marker NP written production test.** Participants are then tested on their ability to produce one-marker NP descriptions. On each trial, participants see an image and are required to type in the corresponding NP(s). Verb forms are provided for Case Only and Number & Case trials. Feedback is provided (an (in)correct-answer sound is played, along with the image and correct description). There are 16 trials total (randomised): 4 trials for each of the types they have been trained on so far.

**Phase 5, two-marker NP production tests.** In the two critical testing blocks, participants must provide first written, then oral descriptions which include the held-out phrase type: two marker NPs, with plural patients (Figure 1, bottom row). The written production task is identical to Phase 4, except it only includes the held-out trial types (12 trials, 3×4 events randomly chosen) and no feedback is given. This written task is added with the purpose of familiarising participants with the held-out trial types prior to the final oral production test phase and will not be included in our analyses.

Finally, participants are asked to produce oral descriptions for *all* trial types in the language. On each trial, participants see an image and are asked to provide a description aloud. As in the previous written production trials, participants are provided with the corresponding verb form when necessary. Feedback is provided (as described above) *only* when the target description does not contain a two-marker NP. There are 58 trials total (randomised): 36 two-marker NP trials, 6 trials of each of the three one-marker NP trial types, and four bare noun trials.

## Results

Recall that, based on Universal 39 (Greenberg, 1963), participants are predicted to produce number markers closer to the noun stem than case markers. This should hold for both the pre- and post-nominal conditions. Our working hypothesis is that these orders are preferred because they reflect the scopal relations among morphemes. Figure 2 is a stacked histogram, showing the percentage of participants whose oral productions follow scope in 0-100% of trials across both conditions. Experiment 1 results (with English speakers) are on the left-hand side. For critical trials, 95% of participants are (almost) perfectly consistent, producing two-marker NPs in the predicted order 95-100% of the time. We ran a logistic mixed-effects regression model predicting use of scope-isomorphic morpheme orders on two-marker NPs dur-

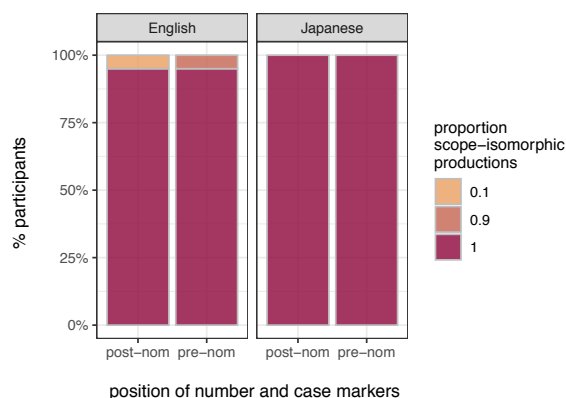


Figure 2: Percentage of participants in Experiments 1 (English) and 2 (Japanese) who produced scope-isomorphic responses a given proportion of the time (rounded to one decimal), ranging from 0% of the time (yellow) to 100% of the time (dark red). Results are split by Marker Position (pre- vs. post-nominal).

Table 1: Model output for Experiment 1.

	$\beta$	$SE$	$z$	$Pr(>  z )$
(Intercept)	13.398	3.213	4.169	< 0.001
Marker Position	-0.219	2.428	-0.090	0.928

ing oral production by Marker Position (pre-nominal vs. post-nominal).<sup>3</sup> As shown in Table 1, the intercept (grand mean of scope-isomorphic productions across participants in both conditions) is positive and significant, confirming that the average proportion of scope-isomorphic productions ( $P \approx 1$ ) is above chance. The effect of Marker Position is not significant, confirming that this preference holds regardless of the pre- or post-nominal positioning of the markers.

## Experiment 2

The results of Experiment 1 are consistent with the hypothesis that scope relations—here between number and case morphemes—determine proximity to the noun stem. Importantly, we can rule out the effect of raw or bigram frequency in driving our results, since these were held constant in our stimuli. However, an alternative explanation is that our result reflects the fact that English overtly marks (plural) number but it does not have morphological case marking (aside from perhaps the genitive). Exactly how this would lead to a preference for placing number closer than case is not totally clear. Perhaps familiarity with, or accessibility of the number marker leads English speakers to place it closer to the

<sup>3</sup>In all models, fixed effects were sum coded unless stated otherwise, and random intercepts for both items (noun) and participants were included. The DV consists of a binary variable marking the presence and absence of scope-isomorphism in each oral production trial (1 for a scope-isomorphic pattern, 0 for an anti-scopal pattern).

noun. To rule this out, we replicated Experiment 1 with native speakers of Japanese. In contrast to English, Japanese overtly marks cases (including accusative) via suffixation; however, the marking of plurality is exceptional (Nakanishi & Tomioka, 2004). The closest thing to number marking *on nouns* are the associative plural classifiers or collectivising suffixes (*-kata*, *-tachi*, *-ra*, *-domo*). Number is typically expressed instead via plural words (which appear after the case inflected noun), reduplication or numeral words (which precede the noun). Japanese speakers should therefore have no trouble acquiring a novel accusative case marker, and if anything should find the case marker more familiar/accessible than the number marker.

## Methods

Experiment 2 is identical to Experiment 1, with one difference: the input lexicon. Rather than using a language with English-like phonotactics, the lexicon for Experiment 2 matched Japanese phonotactics. The preregistered design and analysis plan for Experiment 2 is accessible at [osf.io/akcyp](https://osf.io/akcyp).

**Participants** Forty native Japanese speakers were recruited from Waseda University’s student database. Participants were paid ¥1000 for a 35-min-long experimental session. Note that all participants spoke English as an L2.

**Input language** Lexical items in the language were displayed in Katakana (instead of Latin) script. The three semi-nonce verbs (which contain the stem of the existing verbs in Japanese) are: ケルラ ([ke<sup>+</sup>rura]), ナグラ ([na<sup>+</sup>gura]) and サスラ ([sa<sup>+</sup>sura]), which refer to ‘kicking’, ‘punching’ and ‘pointing’ respectively. The (trisyllabic) nonce nouns are: ソギナ ([sogi<sup>+</sup>na]), ダクメ ([daku<sup>+</sup>me]), ネチビ ([ne<sup>+</sup>ɕibi]), and タソヌ ([taso<sup>+</sup>nu]), naming four characters (a burglar, a chef, a cowboy, and a waitress). The two nonce markers (one for number, one for case) are randomly chosen from the following set: セヒ ([se<sup>+</sup>hi]), ギト ([gi<sup>+</sup>to]), ヨザ ([yo<sup>+</sup>za]). Word order in sentences was Verb-Agent-Patient. Half of the participants were assigned to each of two conditions as per Experiment 1 (i.e. pre-nominal or post-nominal morphology). Auditory stimuli were recorded in a sound-attenuated room by a 28yo female speaker of Japanese.

**Procedure** The experiment was conducted in a quiet room, with all instructions provided in Japanese, and a Japanese-speaking experimenter. Participants were told that they would be learning part of a foreign language. The session proceeded exactly as outlined for Experiment 1.

## Results

The proportion of participants whose oral productions follow scope in 0-100% of trials are shown in Figure 2. The results from Experiment 2 are on the right-hand side. All participants produced number consistently (95-100%) closer to the noun than case. This was true in both the pre-nominal or post-nominal marker conditions. We ran a logistic mixed-effects

Table 2: Model output comparing Experiment 1 and 2.

	$\beta$	$SE$	$z$	$Pr(>  z )$
(Intercept)	12.112	1.966	6.160	< 0.001
Marker Position	-0.0295	1.302	-0.227	0.821
Experiment	-0.012	1.303	-0.009	0.993
Marker Position $\times$ Experiment	0.05	1.302	0.038	0.970

model predicting scope-isomorphic productions by Marker Position (pre- vs. post-) and Experiment (Japanese vs. English). As shown in Table 2, the intercept is positive and significant, confirming that the proportion of scope-isomorphic productions is above chance. The non-significant effects of Marker Position and Experiment confirm that this preference holds regardless of pre- or post-nominal positioning of the markers, and regardless of the native language of participants.

### Experiment 3

Experiments 1 and 2 demonstrate that learners have a natural preference to produce number morphology closer to the noun stem than case. These results hold for pre- and post-nominal orders, suggesting that the preference is not driven by linear order: number appears before case in post-nominal orders, but after case in pre-nominal orders. Our results hold for speakers of both English and Japanese, suggesting that they are not driven by L1 knowledge: familiarity with a particular morpheme (number or case respectively) does not mean it is placed closer to the stem. Frequency cannot explain the preference either: markers for case and number occur with equal frequency, as does each stem+morpheme bigram. The parsability of the morphemes is also the same, since frequencies of stem+morpheme forms relative to stems alone is the same for each. We thus conclude that the results obtained so far are consistent with a bias towards scope-isomorphism.

While our results suggest the bias is very strong (almost all participants uniformly preferred scope-isomorphic orders), in natural language, competing pressures may be present. One such pressure, prominent in models of morphological learning comes from the notion of locality. Dependencies between morphemes (e.g., between an allomorph and the stem that triggers it) tend to be local, or adjacent (Embick, 2010; Moskal, 2015; Bobaljik, 2012). In Experiment 3, we test the strength of the scope-isomorphic bias in the face of a competing locality bias. To do this, we use contextual allomorphy: the form of the case marker is dependent on the lexical and phonological identity of the noun. Because this creates a dependency between the noun stem and the case marker, a locality bias would predict that these two elements should be adjacent. The effect of the scope-isomorphism bias uncovered in Experiments 1 and 2 may override the effect of a locality bias. Alternatively, the locality bias may interfere with the placement of number in closer proximity to the noun stem, leading to a higher proportion of anti-scopal order productions (typologically rare) in the presence of stem-dependent case allomorphy.

## Methods

**Participants** Forty-four English speakers were recruited and compensated as for Experiment 1. They were evenly divided between four conditions, as described below. Following our exclusion criteria, the data of four participants were excluded from analysis.

### Input languages

This was a 2x2 design, with Marker position (pre- and post-) and Allomorphy (no allomorphy vs. case allomorphy) varying between-subjects. The input language in no allomorphy conditions was as in Experiment 1, except that case and number markers appeared as bound morphemes (i.e., affixes) on the noun when presented in text form (no spaces). The input language in the case allomorphy conditions differed additionally in having *two* accusative case markers, which alternated based on the length of the noun: one marker appeared with bisyllabic nouns ('negid', 'tumbat'), the other with monosyllabic nouns ('vaem', 'nork').

**Procedure** The procedure was identical to Experiment 1, except that in two-marker written trials, participants could not advance to the next trial until they typed the correct number of characters. This encouraged participants to produce both two markers together.

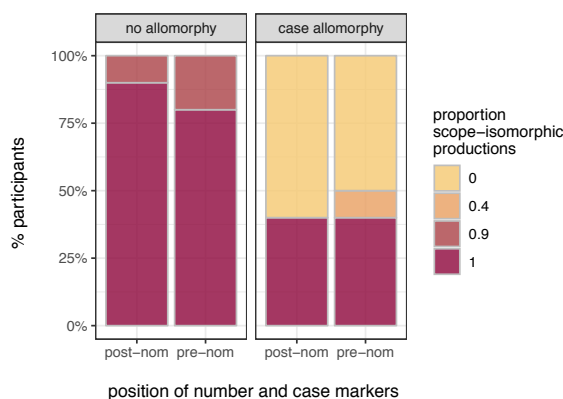


Figure 3: Percentage of participants in Experiment 3 who produced scope-isomorphic responses a given proportion of the time, ranging from 0% of the time (yellow) to 100% of the time (dark red). Results are split by Marker Position (pre- vs. post-nominal) and Allomorphy (no allomorphy vs. case allomorphy).

## Results

Figure 3 shows the percentage of participants whose oral productions follow scope in 0-100% of trials across all four conditions. For the no allomorphy conditions, we replicate our previous findings: participants strongly prefer the scope-isomorphic order, with the number marker closer to the noun than case. By contrast, in the case allomorphy conditions,

Table 3: Model output for Experiment 3

	$\beta$	<i>SE</i>	<i>z</i>	<i>Pr(&gt;  z )</i>
(Intercept)	15.148	4.557	3.324	< 0.001
Marker Position	0.381	4.355	0.087	0.930
Allomorphy	-27.506	5.386	-5.107	< 0.001
Marker Position $\times$ Allomorphy	-0.529	4.691	-0.113	0.910

this pattern is reversed, with most participants producing case closer to the noun. This was confirmed by a logistic mixed-effects regression model predicting use of scope-isomorphic order by Marker Position, and Allomorphy<sup>4</sup>. As shown in Table 3, there is a significant drop in the use of scope-isomorphic orders in the case allomorphy condition.

## Discussion

In the experiments reported here, speakers are trained on a language with distinct number and case morphemes, but the relative order of those morphemes is held out. When required to produce both morphemes together during testing, we found that participants' default inference is to place number closer to the noun stem than case (regardless of whether the markers were pre- or post-nominal). This bias provides a potential causal link between human cognition, and a typological generalisation known as Universal 39 (Greenberg, 1963). Importantly, we found strong evidence for this bias across two populations which differ in terms of their prior experience with case and number markers; English marks number but not case, while Japanese marks case but not number. This suggests our results cannot be explained by relative familiarity with these markers. Furthermore, the observed preference is not dependent on distributional information in the input: case and number markers never appear together, and have the same frequency during training. We have suggested that this bias is driven by scope relations among the markers. In particular, case (which marks the grammatical role of the noun in the event) scopes higher than number (which modifies the set properties of the entity), and linear proximity should reflect scope (Bybee, 1985; Rice, 2000; Culbertson & Adger, 2014). While this order is inferred by default, results from Experiment 3 revealed that the presence of stem-dependent contextual allomorphy for case led many participants to place the case morpheme closer to the conditioning noun. This suggests that the default preference may interact with other constraints—i.e., imposed by morphophonological rather than semantic dependency relationships—as predicted by theories of locality (e.g., White et al., 2018; Embick, 2010). Whether such allomorphy patterns are sensitive to locality in natural language points to the need for additional typological research (although see Moskal, 2015).

<sup>4</sup>The fixed effect of Allomorphy was treatment coded (instead of sum coded) so we could directly compare case allomorphy to the baseline no allomorphy.

## Conclusion

Our results show that in the absence of explicit evidence, language learners default to a typologically common order of morphemes: with number more proximal to the noun stem than case. This supports a hypothesised link between human cognition and Greenberg's Universal 39. However, this observed bias in principle interacts with constraints on locality driven by morphophonological dependencies.

## Data accessibility

The data that support the findings of this study are openly available in the Open Science Foundation repository at <https://doi.org/10.17605/OSF.IO/9FA3V>.

## Acknowledgements

We thank Maki Kubota and Alexander Martin for helping in the recording of the experimental stimuli.

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 757643, held by JC).

## References

- Baayen, H. (1993). On frequency, transparency and productivity. In *Yearbook of morphology 1992* (pp. 181–208). Springer.
- Baker, M. (1985). The Mirror Principle and Morphosyntactic Explanation. *Linguistic Inquiry*, 16(3), 373–415. Retrieved from [www.jstor.org/stable/4178442](http://www.jstor.org/stable/4178442)
- Bobaljik, J. D. (2012). *Universals in comparative morphology: Suppletion, superlatives, and the structure of words* (Vol. 50). Cambridge, MA: MIT Press.
- Bouchard, D. (2002). *Adjectives, number and interfaces: Why languages vary*. Amsterdam: Elsevier. doi: 10.1353/lan.2006.0187
- Bybee, J. L. (1985). *Morphology: A study of the relation between meaning and form* (Vol. 9). Amsterdam: John Benjamins Publishing.
- Cinque, G. (2005). Deriving Greenberg's Universal 20 and its exceptions. *Linguistic Inquiry*, 36(3), 315–332. doi: 10.1162/0024389054396917
- Culbertson, J., & Adger, D. (2014). Language learners privilege structured meaning over surface frequency. *Proceedings of the National Academy of Sciences*, 111(16), 5842–5847. doi: 10.1073/pnas.1320525111
- Culbertson, J., Smolensky, P., & Legendre, G. (2012). Learning biases predict a word order universal. *Cognition*, 122(3), 306–329. doi: 10.1016/j.cognition.2011.10.017
- Embick, D. (2010). *Localism versus globalism in morphology and phonology*. Cambridge, MA: MIT Press.
- Fedzechkina, M., Chu, B., & Jaeger, T. F. (2018). Human information processing shapes language change. *Psychological science*, 29(1), 72–82. doi: 10.1177/0956797617728726

- Fedzechkina, M., Jaeger, T. F., & Newport, E. L. (2012). Language learners restructure their input to facilitate efficient communication. *Proceedings of the National Academy of Sciences*, 109(44), 17897–17902. doi: 10.1073/pnas.1215776109
- Futrell, R., Mahowald, K., & Gibson, E. (2015). Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences*, 112(33), 10336–10341. doi: 10.1073/pnas.1502134112
- Givón, T. (1979). *On understanding grammar*. New York: Academic Press.
- Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of human language* (pp. 73–113). Cambridge, MA: MIT press.
- Grimshaw, J. (1986). A morphosyntactic explanation for the mirror principle. *Linguistic Inquiry*, 17(4), 745–749. Retrieved from <http://www.jstor.org/stable/4178514>
- Hay, J. (2001). Lexical frequency in morphology: Is everything relative? *Linguistics*, 39(6), 1041–1070. doi: 10.1515/ling.2001.041
- Hay, J., & Plag, I. (2004). What constrains possible suffix combinations? on the interaction of grammatical and processing restrictions in derivational morphology. *Natural Language & Linguistic Theory*, 22(3), 565–596. doi: 10.1023/B:NALA.0000027679.63308.89
- Hupp, J. M., Sloutsky, V. M., & Culicover, P. W. (2009). Evidence for a domain-general mechanism underlying the suffixation preference in language. *Language and Cognitive Processes*, 24(6), 876–909. doi: 10.1080/01690960902719267
- Manova, S., & Aronoff, M. (2010). Modeling affix order. *Morphology*, 20(1), 109–131. doi: 10.1007/s11525-010-9153-6
- Moskal, B. (2015). Limits on allomorphy: A case study in nominal suppletion. *Linguistic Inquiry*, 46(2), 363–376. doi: 10.1162/LING\_a\_00185
- Nakanishi, K., & Tomioka, S. (2004). Japanese plurals are exceptional. *Journal of East Asian Linguistics*, 13(2), 113–140. doi: 10.1023/B:JEAL.0000019058.46668.c1
- Plag, I., & Baayen, H. (2009). Suffix ordering and morphological processing. *Language*, 109–152. doi: 10.1353/lan.0.0087
- Rice, K. (2000). *Morpheme order and semantic scope: Word formation in the Athapaskan verb* (Vol. 90). Cambridge, UK: Cambridge University Press.
- Rice, K. (2011). Principles of affix ordering: An overview. *Word Structure*, 4(2), 169–200. doi: 10.3366/word.2011.0009
- Ryan, K. (2010). Variable affix order: grammar and learning. *Language*, 86(4), 758–791. doi: 10.1353/lan.2010.0032
- Tabullo, Á., Arismendi, M., Wainseboim, A., Primero, G., Vernis, S., Segura, E., ... Yorio, A. (2012). On the learnability of frequent and infrequent word orders: An artificial language learning study. *The Quarterly Journal of Experimental Psychology*, 65(9), 1848–1863. doi: 10.1080/17470218.2012.677848
- White, J., Kager, R., Linzen, T., Markopoulos, G., Martin, A., Nevins, A., ... van de Vijver, R. (2018). Preference for locality is affected by the prefix/suffix asymmetry: Evidence from artificial language learning. In *Proceedings of NELS 48* (p. 207–220). Amherst, MA: Graduate Linguistics Student Association.
- Wilson, C. (2003). Experimental investigation of phonological naturalness. In *Proceedings of WCCFL 22* (Vol. 22, pp. 533–546).
- Wunderlich, D. (1993). Funktionale kategorien im lexikon. In F. Beckmann & G. Heyer (Eds.), *Theorie und praxis des lexikons* (pp. 54–73). Berlin: Walter de Gruyter.