

Learners are sensitive to hierarchical structure in noun phrases containing demonstratives but not numerals

Alexander Martin (University of Edinburgh), **Klaus Abels** (University College London), **David Adger** (Queen Mary University London), **Jennifer Culbertson** (University of Edinburgh)

Previous research has argued that learners infer word order patterns in a new language based on knowledge about underlying structure, rather than linear order (Culbertson & Adger, 2014; Martin, Ratitamkul, Abels, Adger, & Culbertson, in press). In a series of experiments, across multiple speaker populations, participants were taught a version of their native language in which the order of a single modifier (adjective, numeral or demonstrative) relative to the noun was swapped (e.g., English speakers learned that “green car” was produced “car green” in the new language). Participants then had to guess the relative order of multiple modifiers (e.g., “two green cars”). Across experiments, participants inferred orders which reflected not their native language order transposed (e.g., “cars two green”), but instead the underlying hierarchical structure in this domain: adjectives closest to the noun, then numerals, then demonstratives (Cinque, 2005; Abels & Neeleman, 2012, e.g., “cars green two”). This preference is also reflected typologically, with orders that transparently reflect this hierarchical structure (termed *homomorphic* orders) overwhelmingly outnumbering orders that do not (this asymmetry is commonly referred to as Universal 20, Greenberg, 1963).

However, these previous studies employed a highly non-naturalistic design, relying on native-language materials, presented orthographically. In this context, a conscious strategy of flipping—reversing the order of the words to determine the new order—could explain why participants chose homomorphic orders, without tapping into hierarchical language structure at all. In the present study, we reexamine these effects using a more naturalistic experimental design that is not based on native-language material. We designed naturalistic “foreign” languages which we taught to English-speakers recruited through Amazon’s Mechanical Turk. Participants were instructed that they would be learning part of a new language called Nápijò, spoken by around 10,000 people in a rural region of Southeast Asia. In actuality, Nápijò was an artificial language composed of three nouns and a series of modifiers.

In Experiment 1, Nápijò included the nouns /e'je/ (*feather*), /u'hu/ (*ball*), and /i'ti/ (*mug*), two adjectives (/pu'ku/, *red* and /ta'ka/, *black*), and two additional items /ho'no/ and /hi'mi/ which served as either the demonstratives *this* and *that* or the numerals *two* and *three*, depending on the condition participants were assigned to. That is, participants learned adjectives and *either* numerals *or* demonstratives. Participants were taught word meanings by hearing the labels presented with simple cartoon scenes. Objects were depicted on a table standing in front of a girl. Participants first learned the object names by seeing greyscale images of the objects. To learn the meanings of the modifiers, participants saw a modified object (e.g., a red feather, two mugs)¹, and heard the noun label *followed* by the modifier label. That is, in Nápijò, all modifiers were postnominal. This means that participants could not rely on their English knowledge to learn Nápijò word order.

Once participants had learned the individual word meanings, and how nouns combined with single modifiers in the language, they moved on to a testing phase where they were presented with visual stimuli that involved a colour *and* either a numeral (e.g., a group of three red feathers) or a demonstrative (e.g., the cartoon girl pointing to a red feather in front of her). They were asked to choose how to describe the scene in Nápijò and were given two options, one with a homomorphic order, and one with a non-homomorphic order. Participants’ homomorphic preferences in Experi-

¹Demonstrative meanings were taught by showing the girl pointing at a proximal or a distal object.

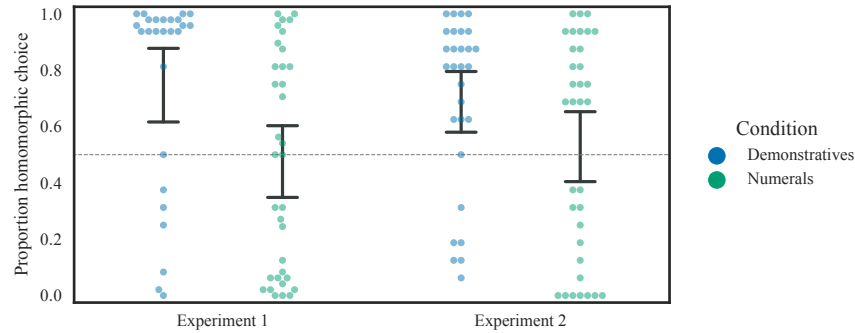


Figure 1: Proportion homomorphic preference in each experiment by condition. Each point represents an individual participant and error bars represent 95% confidence intervals.

ment 1 are plotted on the left side of fig. 1. We found a homomorphism preference in the Demonstratives condition ($\chi^2(1) = 11.35, p < 0.001$) but not in the Numerals condition ($\chi^2(1) < 1$).

In Experiment 2, we explored whether the preference for homomorphic orders could be amplified by a more naturalistic distribution of modifier types. In natural languages (and in the previous experiments discussed above) there are typically many more adjectives than numerals, and relatively few demonstratives. This may serve to distinguish the categories more clearly to learners. In this version, Nápijò adjectives were /ta'kas/ (*black*), /pu'kuf/ (*red*), /ka'paθ/ (*blue*), and /ku'tuf/ (*green*). Participants' homomorphic preferences in Experiment 2 are plotted on the right side of fig. 1. Again, we found a homomorphism preference in the Demonstratives condition ($\chi^2(1) = 10.37, p < 0.01$) but not in the Numerals condition ($\chi^2(1) < 1$).

To summarise, in both Experiments 1 and 2, we found a clear homomorphism preference for demonstratives and adjectives. This confirms that in a more naturalistic task, where participants are unlikely to use an explicit strategy of flipping English words to determine order in the new language, a homomorphism preference is still found. However, we did not find such a preference for numerals and adjectives, regardless of the size of the adjective class. Interestingly, both Culbertson and Adger and Martin et al. point to numerical trends in their data suggesting this same asymmetry. Indeed, non-homomorphism between numerals and adjectives, or numerals and demonstratives is in fact more common cross-linguistically (35 and 64 languages respectively) than non-homomorphism between adjectives and demonstratives (27 languages) (numbers calculated from Dryer, 2018). This may reflect the fact that adjectives and demonstratives are more distant from one another in terms of underlying hierarchical structure. While our results support this claim, an additional experiment testing homomorphic preferences for demonstratives and numerals would need to be carried out. If our claim is true, then we would expect those preferences to pattern with the numeral/adjective condition of our experiments. Discussion will focus on possible underlying causes of the demonstrative/numeral asymmetry.

References

- Abels, K. & Neeleman, A. (2012). Linear asymmetries and the LCA. *Syntax*, 15(1), 25–74.
- Cinque, G. (2005). Deriving Greenberg's Universal 20 and Its Exceptions. *Linguistic Inquiry*, 36(3), 315–332. JSTOR: 4179327
- Culbertson, J. & Adger, D. (2014). Language learners privilege structured meaning over surface frequency. *PNAS*, 111(16), 5842–5847.
- Dryer, M. S. (2018). On the Order of Demonstrative, Numeral, Adjective and Noun. *Language*.
- Greenberg, J. H. (1963). *Universals of Language*. Cambridge: MIT Press.
- Martin, A., Ratitamkul, T., Abels, K., Adger, D., & Culbertson, J. (in press). Cross-linguistic evidence for cognitive universals in the noun phrase. *Linguistics Vanguard*.