

Semantic influences on emergent preferences of word order: Evidence from silent gesture

Jida Jaffan

Dept. of Linguistics
University of Toronto
jida.jaffan@mail.utoronto.ca

Gabrielle Klassen

Dept. of Spanish & Portuguese
University of Toronto

Jordan Ziqi Yang

Dept. of Linguistics
University of Toronto

Daphna Heller

Dept. of Linguistics
University of Toronto

daphna.heller@utoronto.ca

Abstract

Across the world's languages, some word orders are more common. We focus on noun phrases, where it is more common for adjectives to follow the noun than to precede it. Because the interpretation of adjectives depends on the noun they modify, we propose and evaluate the new hypothesis that the order N-ADJ is more prevalent because it is beneficial for semantic processing. In a silent gesture task, speakers of four typologically-unrelated languages (English, Mandarin, Arabic and Spanish) communicated noun phrase meanings to a partner. We find, first, that our task tracks the typologically-preferred orders of nouns, adjectives and numerals in the noun phrase. More importantly, we find support for our semantic processing hypothesis: size adjectives, whose interpretation depend more on the noun they modify, were more likely to be gestured after the noun than shape adjectives whose interpretation is less dependent on the noun they modify.

Keywords: word order; noun phrase; adjectives; silent gesture; linguistic universals; cognitive biases

Introduction

Typological research reveals that some word orders are more common than others. In the noun phrase, for example, *adjectives* more frequently follow the noun (N-ADJ: 64%) than precede it (ADJ-N: 27%), with a small minority of languages (8%) exhibiting a variable order (Dryer, 2013). In comparison, the location of the *numeral* relative to the noun is more balanced: the numeral follows the noun (N-NUM) in 52.6% of the world's languages, and precedes the noun (NUM-N) in 41.5% of languages, with again a small minority (5.6%) exhibiting a variable order (Dryer, 2013). Interestingly, further asymmetries are observed in the combination of these elements: languages where the adjective and the numeral occur on the same side of the noun – known as *harmonic* orders – are more common (N initial: 52%; N final: 27%) – see in Table 1¹.

Table 1: Typological distribution of word order in the noun phrase. Harmonic orders are marked in grey.

	ADJ-N	N-ADJ
N-NUM	4%	52%
NUM-N	27%	17%

¹ Note that the overall values reported in this paragraph are somewhat different from the marginal means in Table 1: this arises

Prior work has focused on the *relative* order of elements. In his seminal work on language universals, Greenberg (1963) includes Universal 18: “When the descriptive adjective precedes the noun, the demonstrative, and the numeral, with overwhelmingly more than chance frequency, does likewise.” In other words, there is a preference for harmonic orders, especially when the adjectives precede the noun.

Previous psycholinguistic work has addressed the issues of harmony in noun phrases. For example, Culbertson and Newport (2017) demonstrated that harmonic patterns are preferred in learning. Specifically, children and adults were taught an artificial grammar where modifiers preceded the noun 70% of the time. And while adults subsequently produced noun phrases that were similar to the input, children altered the input to create harmonic orders. These patterns provided evidence for the cognitive underpinning of the typological patterns. Building on Greenberg's (1963) universals, Culbertson and Newport (2017) proposed the principle of harmony, which reflects a preference for a single rule over multiple rules (e.g., “modifiers should follow the noun” or “modifiers precede the noun”).

Schouwstra, Kirby, and Culbertson (2017) further examine the emergent order in noun phrases using a silent gesture study, where target expressions included a noun, an adjective, a numeral, and a demonstrative (e.g., *these four striped triangles* or *those three dotted squares*). They found, first, that gestured adjectives overwhelmingly followed the noun, despite the fact that in English, the native language of their participants, adjectives generally precede the noun. Furthermore, they found that the order of modifiers was guided by their semantic scope: adjectives appeared closest to the noun, then numerals (which count the noun-adjective combination), and finally determiners (which pick out the noun-adjective-numeral combination).

Our goal here is to focus on the typological preference for the adjectives to follow the noun (i.e., the prevalence of N-ADJ order). We propose the novel hypothesis that this order is preferred because it is beneficial for **semantic processing**. This proposal arises from combining what we know about language processing with what we know about the interpretation of adjectives. First, it is well-established in the psycholinguistics literature that the processing of noun phrase – and language more generally is incremental (Eberhard, Spivey-Knowlton, Sedivy, and Tanenhaus, 1995;

from a difference in how languages with a variable word order are counted.

Sedivy, Tanenhaus, Chambers, & Carlson, 1999). Second, it is well-established in the semantics literature that the interpretation of adjectives depends on which noun they modify (Kamp & Partee, 1995; Kennedy & McNally, 2005; Toledo & Sassoon, 2011). For example, some adjectives are *intersective*, whereas other adjectives are *subsective* (Kamp & Partee, 1995). When an intersective adjective combines with a noun (e.g., *a square hairbrush*), the meaning of the noun phrase is derived using intersection: the set of objects denoted by *square* is intersected with the set of objects denoted by *hairbrush*, and the noun phrase denotes an object that is both square and a hairbrush. By contrast, when a subsective adjective combines with a noun (e.g., *a big hairbrush*), the adjective and the noun do not have similar contributions to the meaning; instead, the interpretation is derived by subsecting the set of hairbrushes. In other words, the interpretation of “*big*” – and other scalar adjectives – depends on a standard that is set by the noun they modify.

Taken together, this predicts a *processing advantage* for a noun phrase where the adjective follows the noun, compared to processing the adjective before the noun becomes available in the speech stream. Importantly, we are not assuming that adjectives are processed only after the noun (cf. Kamp & Partee, 1995): it is well established that adjectives are processed incrementally even when they precede the noun (Eberhard, et al., 1995; Sedivy, et al., 1999). Instead, the idea is that the order N-ADJ is more widely found because it gives rise to a semantic processing benefit.

We reason that if there is a cognitive bias for the order N-ADJ, this bias may surface when individuals tailor a message that is not subject to pre-existing linguistic conventions. Therefore, we tested this hypothesis using a silent gesture task. This method has been successfully used to study word order biases both at the sentence level – the ordering of Subject, Object, and Verb (Goldin-Meadow, So, Özyürek, & Mylander, 2008; Gibson et al., 2013), and at the level of the noun phrase (Schouwstra et al., 2017).

Our method differed slightly from these studies, as our participants were not asked to tailor a message for an imagined addressee, but instead were communicating with a lab confederate. In this setup, the participant had to get the confederate to choose one image out of an array of four – see Figure 1.

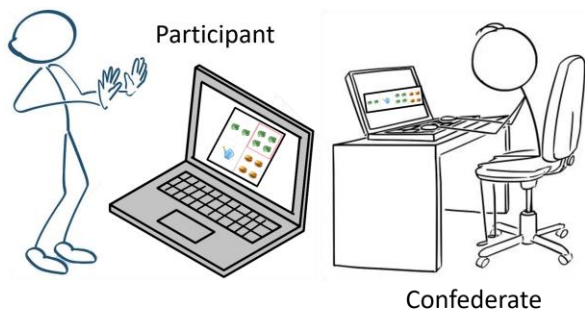


Figure 1: The communicative setup

One reason for using a lab confederate – rather than having participant gesture to the experimenter as in previous studies – was to ensure that participants had no assumptions of their partner’s knowledge, and would therefore gesture all the information necessary to identify the target (which varied across conditions – more below). A second reason for using a confederate was to avoid a situation where a naïve addressee would try to be cooperative and contribute their own gestures or guess the target based on partial information (e.g., if only the adjective was gestured, and not the noun).

We include four conditions – see Figure 2:

- (a) ADJ-SIZE, which is intended to elicit a noun and a size adjective;
- (b) ADJ-SHAPE, which is intended to elicit a noun and a shape adjective;
- (c) NUMERAL, which is intended to elicit a noun and a numeral;
- (d) NUM-ADJ, which is intended to elicit three elements: a noun, an adjective and a numeral. To simplify the displays, all adjectives elicited in this condition were shape adjective.

Note that, across all conditions, the elements we elicited depended not just on properties of the target image itself, but also on the properties of other images in the display.



Figure 2: Example displays for the four conditions: the image to be communicated is marked in red.

First, if considerations of semantic processing play a role in guiding the ordering of elements, we would expect adjectives to follow the noun more in the ADJ-SHAPE condition than in the ADJ-SIZE condition. Furthermore, given that this reasoning does not apply to numerals, the

relative ordering of nouns and numerals are not expected to differ from chance.

A further aspect of our design is the native languages of the speakers tested. We wanted to ensure that any effects of word order are above and beyond effect of the native language; previous work using silent gesture of noun phrases only examined native speakers of English (e.g., Schouwstra et al., 2017). Therefore, we tested native speakers of four typologically-different languages: English, Mandarin, Arabic, and Spanish. English, exemplified in (1), and Mandarin, exemplified in (2) both exhibit the order N-ADJ; both languages also exhibit NUM-N order. Note that speakers of languages that use the less preferred order (such as these) provide a stronger test to whether the typologically preferred order will be the one to emerge in a silent gesture task.

- (1) ENGLISH
 a. The big hairbrush
 b. The square hairbrush
 c. Four hairbrushes
 d. Four square hairbrushes
- (2) MANDARIN
 a. dà shūzi
 big hairbrush
 b. fāngxíng shūzi
 square hairbrush
 c. sì bǎ shūzi
 four [classifier] hairbrush
 d. sì bǎ fāngxíng shūzi
 four [classifier] square hairbrush

Our third language, Arabic exhibits the order N-ADJ (3a-b); here NUM is variable, with the number 1 appearing after the noun (3c), the number 4 appearing before the noun (3d), and the number 3 being in free variation (3e).

- (3) ARABIC
 a. firšaye kbire
 hairbrush big
 b. firšaye mrabaʔa
 hairbrush square
 c. firšaye wahde
 hairbrush one
 d. arbaʔa firaši
 four hairbrush
 e. tlat firaši / il-firaši it-tlate
 three hairbrushes / the-hairbrushes the-three

Our final language Spanish, exemplified in (4), is generally N-ADJ, but some adjectives appear in the order ADJ-N. These two syntactic positions are associated with interpretive differences. While characterizing these differences is non-trivial (see e.g., Demonte, 2008), the noun phrase meanings gestured in our task are only appropriate as post-nominal adjectives. Spanish is also NUM-N order.

- (4) SPANISH
 a. cepillo grande
 hairbrush big
 b. cepillo cuadrado
 hairbrush square
 c. cuatro cepillos
 four hairbrushes
 d. cuatro cepillos cuadrados
 four hairbrushes square

Our first question is whether, when producing silent gestures, participants will show a preference for the typologically-preferred orders, as has been previously shown for the order of the elements in the sentences: Subject-Verb-Object (Goldin Meadow, et al., 2008; Gibson et al., 2013). Our second question is whether we can find evidence in favour of our hypothesis that semantic processing plays a role in typological preferences; if so, we would expect size adjectives to follow the noun more because their interpretation is relativized to the noun meaning in a way that shape adjectives are not. Comparing speakers of four typologically different languages should allow us to observe effects that are above and beyond influences of the native language.

Method

Participants

Data from native speakers of Canadian English (n=24) was collected in Toronto, Canada. Data from native speakers of Mandarin Chinese (n=20) was also collected in Toronto, Canada; at the time of testing, participants lived in Canada for three years or less. Data from Spanish speakers (n=25) was collected in Madrid, Spain. Data from Arabic speakers (n=20) was collected in Lebanon and Syria. For all languages, the majority of participants were university students, and each participant was paid \$10 or the locally-appropriate equivalent. Speakers were chosen such that the language of testing was their dominant language; none of the participants had any knowledge of sign language.

Materials and Design

Thirty-two experimental trials were created. Each experimental display contained four images: these were arranged in a 2x2 grid for the participant, and a 1x4 grid for the confederate – see again Figure 1. The images were presented in a different configuration to exclude the possibility of referring to the location of the target image rather than its properties. Across items, the location of the target image in the grid was systematically varied.

Four versions of each display were created – see again Figure 2. In the ADJ-SHAPE condition, the target image contrasted in shape with a second object from the same nominal category (e.g., target: square brush; contrast: round brush). The display also included a second square item, in order to prevent a situation where shape information would

be sufficient to identify the target and hence noun information would not be gestured. In the ADJ-SIZE condition, the target image contrasted in size with the second object from the same nominal category (e.g., target: small clock; contrast: big clock); again, these displays included a second object that was the same size as the target to ensure that both adjective and noun information were gestured. In the NUMERAL condition, the target image contrasted in number: the target was a group of 1, 3, or 4 objects and there was another group of 1-4 objects (e.g., 3 tables and 2 tables), the display also included a second group that was the same number as the target. In the NUM-ADJ condition, the target image contrasted with *two* different images, one in the shape and another in number: for example, three square clocks (target), three round clocks (contrast in shape) and two square clocks (contrast in number).

To avoid confusion in gestures between nouns and adjectives (especially shape adjectives), we chose high-frequency nouns that have been previously shown to be gestured using their function rather than their shape (van Nipsen, van de Sandt-Koenderman, & Krahmer, 2017).

Across participants, each of the experimental displays occurred in all four conditions; however, a single participant only saw a given experimental display in one of the four conditions. In addition to 32 critical trials, each list also included 32 filler trials: these trials were visually similar to the critical trials in that they contained two objects of the same type, but here the target was *not* a member of this pair, but rather one of the other items. This was done in order to ensure that there was no expectation that noun phrases would have a certain form. Thus, each participant saw 64 trials, presented in a random order (the first three were always fillers).

Procedure

Each experimental session was carried out entirely in the relevant language. Participants were told that they would need to communicate an object out of an array to a second person, using silent gestures.

The participant and their partner were seated across a table, each with their own computer – see Figure 1. Participants were not told that the person they were interacting with worked for the lab (a different confederate was employed for each language). Importantly, confederates were naïve to the goals of the experiment.

Participants were told that each trial will be a display with four images, with one highlighted. They would then aim to provide a message using silent gesture that would get their partner to click on that object. Participants were instructed not to use any pointing in their gestures. Participants were further instructed that if they don't provide all the necessary information, the confederate will use a particular gesture to tell them to elaborate. Importantly, this was the only gesture confederates were allowed to produce.

After the participant completed all the trials, they repeated the task with verbal instructions (cf. Gibson et al., 2013). Throughout the session, which lasted about an hour, the participant was video and audio recorded.

Results

Verbal trials were transcribed, and then coded for word order. Gesture trials were transcribed for word order directly from the video recording. For example, a watch was sometimes gestured by looking at one's wrist. Size information was usually gestured with expanding and shrinking gestures. Shape information was usually gestured as the circumference of the shape. Number information was usually gestured by raising fingers. If the participant repeated their gestures, we coded the first gesture sequence that resulted in a successful reference to the target object.

Do gesture patterns follow typology?

We first ask whether, when gesturing noun phrase meanings, speakers exhibit a preference for certain word orders over others, specifically those that are common among the world's languages. To address this question, we first calculated, for each participant, how likely they were to gesture the adjective and the numeral *after* the noun. Figure 3 (top panel) plots word order preferences for each participant in the verbal task and Figure 3 (bottom panel) plots the same information for the critical silent gesture task: the x-axis represent the proportion of gesturing the order N-ADJ, and the y-axis represents the proportion of gesturing the order N-NUM (each data point is one participant, colour coded for native language).

In the verbal task, speakers of the four languages generally exhibit the pattern expected from the grammar of their language. Speakers of both English (red) and Mandarin (blue) generally preferred ADJ-N and NUM-N order². Spanish (yellow) and Arabic (green) speakers produced only N-ADJ order; both also followed the expected order for NUM: which is pre-nominal in Spanish and variable in Arabic.

The gesture data exhibits a different pattern. We note that gesturing gave rise to more variability: this indicates that, when gesturing, speakers did not simply follow the word order dictated by their native language, which opens the possibility of patterns that relate to the typology of noun phrases. Indeed, the patterns observed do not follow the order of the native language. This pattern is strikingly similar to the typological distribution reported in Table 1. First, our participants were most likely to gesture the harmonic order where both the numeral and the adjective follow the noun (our data: 43%, vs. cross-linguistically: 52%). The second most common typological order – the other harmonic order where both the adjective and the numeral precede the noun – was also the second most common here (our data: 28% vs. cross-linguistically: 27%). Interestingly, the non-harmonic order whereby the adjective follows the noun (N-ADJ) but

² The order N-ADJ in English often arose from the production of a relative clause (e.g., “*the clock that's round*”).

the numeral precedes the noun (NUM-N) was more common in our gesture data than it is across the world's language (our data: 27% vs. cross-linguistically: 17%); this may be due to influences from the native languages of our participants. Finally, the second non-harmonic order, whereby the adjective precedes the noun (ADJ-N) but the numeral follows the noun (N-NUM), was strongly dispreferred, as is the case across the world's languages (our data: 2%, vs. cross-linguistically: 4%). The similarity of our gesture data to typological patterns lend further support to the idea that typological patterns arise due to cognitive biases. Specifically, these patterns lend further support to the Culbertson and Newport's (2017) proposal that harmonic orders are preferred over non-harmonic orders.

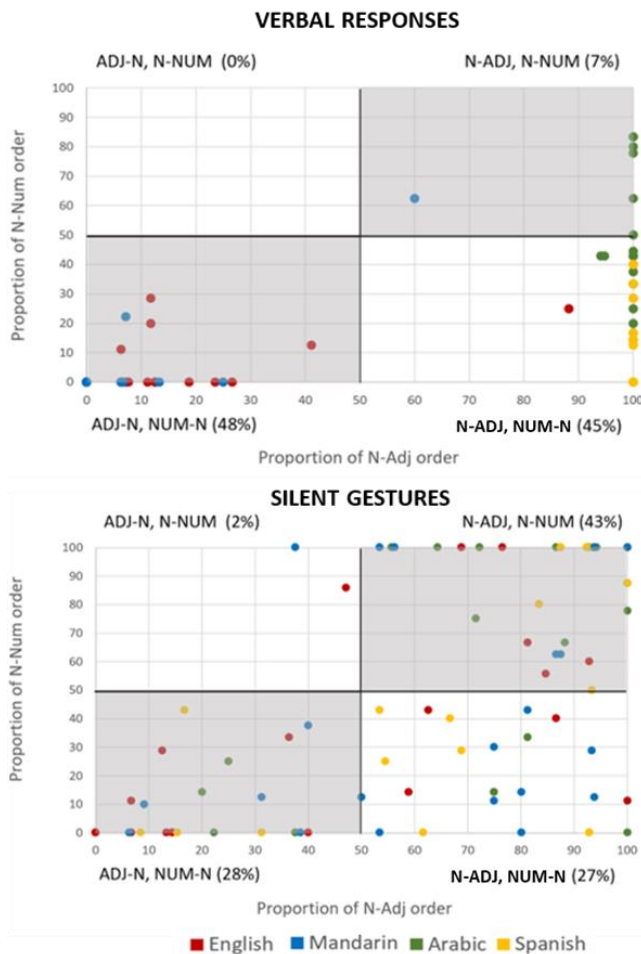


Figure 3. Word order preferences in verbal responses (top panel) and in silent gesture (bottom panel). Each data point represents one speaker, colour-coded by native language.

For the NUM-ADJ condition, we can further ask about the position of the noun compared to the other elements. Figure 4 plots the likelihood of the noun to appear initially, in the middle and at the end of the noun phrase. We note that harmonic order – with the noun either preceding or following both modifiers – are preferred over the case where the noun appears in with one modifier on each side. This is particularly

important for speakers of Arabic and Spanish, both languages that do not exhibit a harmonic order.

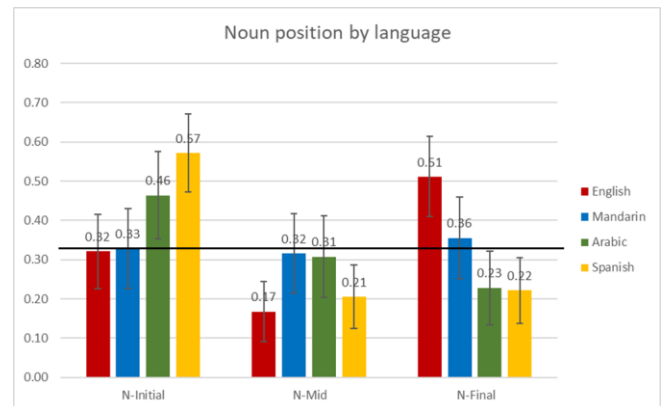


Figure 4. The likelihood of each of the three possible positions of the noun relative to the two modifiers in the NUM-ADJ condition. Chance is .33.

Since gestures for numeral and adjective were used repeatedly, whereas nouns were only used once, one could imagine that the repeated gestures would become conventionalized, and, as a result, speakers will tend to gesture modifiers first, creating a bias for N final. The pattern in Figure 4 suggests that word orders in the silent gesture task were not driven by the asymmetry between old and new gestures.

Effects of adjective semantics

Our second question concerns the origin of the preference for adjectives to follow the noun. Recall that if semantic processing affects the relative order of noun and adjective, we could expect size adjectives to follow the noun more than shape adjectives. We therefore compared the likelihood that the adjective followed the noun in the ADJ-SIZE condition and the ADJ-SHAPE condition, with the NUMERAL condition serving as baseline. These are summarized, for each language, in Figure 5. Qualitatively, we note that across all four languages, the noun was more likely to occur first when followed by a size adjective than a shape adjective; both were more likely to follow the noun than the numeral. Despite these similarities, we also observe an effect of native language order: the likelihood of noun initial was higher in Arabic and Spanish, where adjectives follow the noun, as compared to English and Mandarin, where adjective precede the noun.

These data were analyzed using logistic mixed-effects regression models with random effects for participants (Baayen, Davidson, & Bates, 2008; Jaeger, 2008), as implemented in the *lme4* package of the statistical software R 3.5.1 (Bates, Maechler, Bolker, & Walker, 2015; R Core Team, 2015). The overall model was 3 x 4. The three-level independent variable – condition – was contrast-coded using centered Helmert contrasts. The first coefficient contrasted NUM with ADJ, and the second coefficient contrasted the two ADJ conditions. The four-level independent variable –

language – group the N-ADJ languages and the ADJ-N language, comparing within and across groupings.

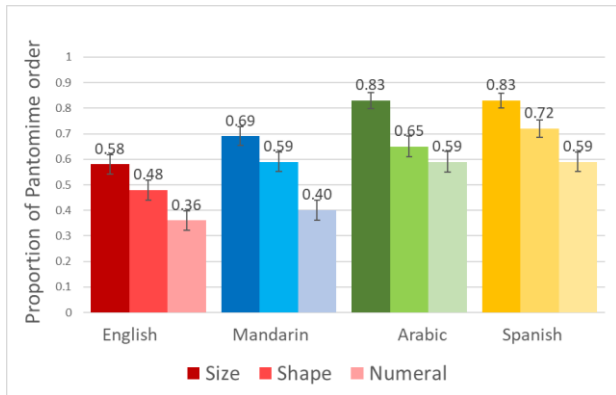


Figure 5: likelihood of N-MOD order, for each of the four languages, in the ADJ-SIZE, ADJ-SHAPE, and NUMERAL conditions.

The model indicated, first, that the noun was more likely to come first when it appeared with an adjective compared to a numeral ($\beta = 7.4089, SE = 1.3861, z = 5.345, p < .0001$). This effect held when examining each language separately, in English ($\beta = 6.4643, SE = 2.5876, z = 2.498, p = .01$), Arabic ($\beta = 12.368, SE = 2.975, z = 4.158, p < .0001$), and Spanish ($\beta = 6.8836, SE = 3.1356, z = 2.195, p = .03$), but was only marginal in Mandarin ($\beta = 4.1075, SE = 2.3664, z = 1.736, p = .08$). This means that adjectives were more likely than numeral to be gestured after the noun, providing support to our hypothesis of semantic processing: the interpretation of adjectives – but not numerals – depends on the noun they modify.

More importantly, the second coefficient revealed that the noun was more likely to appear first when the adjective

encoded size information than when it encoded shape information ($\beta = -0.9683, SE = 0.1550, z = -6.246, p < .0001$). This effect held when analyzing each language separately: English ($\beta = -0.8558, SE = 0.2880, z = -2.972, p = .003$), Mandarin ($\beta = -0.5977, SE = 0.2630, z = -2.272, p = .02$), Arabic ($\beta = -1.492, SE = 0.334, z = -4.466, p < .0001$), and Spanish ($\beta = -0.9525, SE = 0.3520, z = -2.706, p = .006$). This pattern provides strong support to our semantic processing hypothesis that the typological preference for N-ADJ order is driven – at least in part – by having a word order where the adjective information, which depends on the noun, follows it linearly. A further illustration of this pattern is given in Figure 6: here we plot, for each participant, the extent to which they gestured a size adjective after the noun more than a shape adjective: this plot shows that this behaviour is characterized more than half of the participants (46/89), with only 19 participant showing the opposite pattern.

Note that this analysis nevertheless reveals effects of native language: the order N-ADJ was more common for speakers of Arabic and Spanish compared with speakers of English and Mandarin (specifically, our overall model showed that Mandarin and English do not differ, that Arabic and Spanish do not differ, but that Arabic and Mandarin do differ).

General Discussion

This paper aims to examine the cognitive basis for typological preferences regarding word order in the noun phrase. We conducted a silent gesture task with native speakers of four typologically distinct languages that exhibit different word orders within the noun phrase: English, Mandarin, Spanish and Arabic.

First, the relative ordering of noun and adjective and noun and numeral in our task produced a pattern which is strikingly similar to the typological distribution of these orders. Like in the world’s languages, the harmonic orders were more

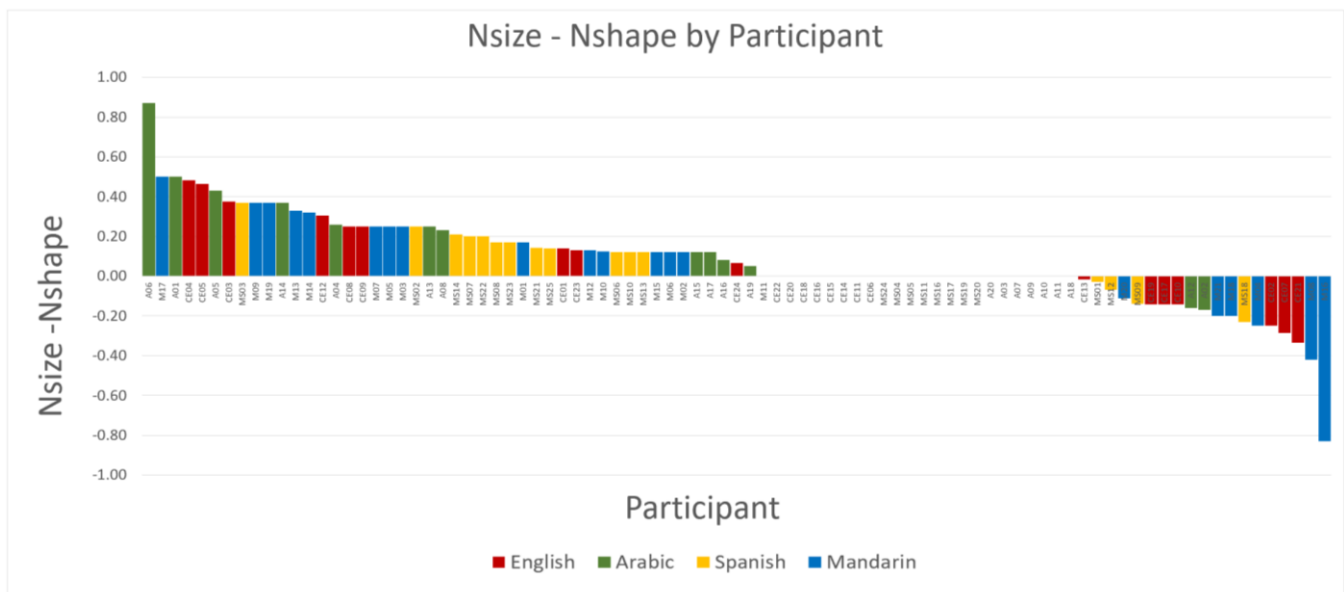


Figure 6. The extent to which each participant was more likely to gesture the size adjective after the noun compared to shape adjective.

common in our gesture task, with most common order – the one where both modifiers follow the noun – being the most commonly gestured. Similarly, the typologically rare order – where the adjective precedes the noun and the numeral follow it – was also rarely gestured by our participants. This finding provides further support to the idea that typology at the macro level arises from cognitive biases at the level of the individual mind (Goldin-Meadow et al., 2008; Culbertson et al., 2012; Gibson et al., 2013; Culbertson & Newport, 2017).

Second, despite the similarity of gestured patterns to typological preferences, we nevertheless observed influences of the native language of participants on the word order used when gesturing. For example, speakers of English and Mandarin were less likely to gesture the order N-ADJ which deviates from their native order, as compared with speakers of Arabic and Spanish, for whom N-ADJ is the order used in language as well (see again Figure 5). Relatedly, English speakers gestured noun final noun-phrases at a level above chance (see again figure 4). These findings are of methodological significance, as they demonstrate the need for a cross-linguistic perspective when conducting experiments that aim to uncover the cognitive underpinnings of linguistic patterns.

Finally, an important goal of this study has been to evaluate a new hypothesis about the typological preference of adjectives to *follow* the noun, testing the hypothesis that this order is beneficial for semantic processing because the interpretation of adjectives depends on the noun they modify. We find that when the adjective encoded information whose interpretation depends on the noun more – specifically, size information – the adjective was more likely to follow the noun more compared with modifier information whose interpretation is less dependent on the noun being modified – specifically, shape information. This finding raises the question of why this effect is not grammaticalized: we may expect languages to place shape information before the noun and size information after. It is notable in this context that some languages, including Spanish and other Romance languages, do exhibit a grammatical pattern that is consistent with this prediction: adjectives that follow the noun receive a restrictive interpretation, whereas when the same adjectives precede the noun, they receive a non-restrictive interpretation (e.g., Demonde, 2008). However, it is likely that this pattern is overridden by the strong preference for harmony (recall that only 8% of the world’s language exhibit a variable order for adjectives).

Taken together, these results contribute to a line of research which takes typological preference at a large scale to arise from cognitive biases at the level of the individual speaker.

Acknowledgments

This work was partially supported by a SSHRC Insight Grant awarded to D. Heller.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for

- subjects and items. *Journal of Memory and Language*, 59, 390–412.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Culbertson, J., & Newport, E. L. (2017). Innovation of word order harmony across development. *Open Mind*, 1(2), 91–100.
- Culbertson, J., Smolensky, P., & Legendre, G. (2012). Learning biases predict a word order universal. *Cognition*, 122(3), 306–329.
- Demonde, V. (2008). Meaning-form correlations and the adjective position in Spanish. In Kennedy, C. & McNally, L. (eds.), *The semantics of adjectives and adverbs*. Oxford University Press.
- Dryer, M. S. (2013). On the six-way word order typology, again. *Studies in Language*. International Journal sponsored by the Foundation “Foundations of Language”, 37(2), 267–301.
- Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of psycholinguistic research*, 24(6), 409–436.
- 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.
- Goldin-Meadow, S., So, W. C., Özyürek, 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.
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59(4), 434–446.
- Kamp, H., & Partee, B. (1995). Prototype theory and compositionality. *Cognition*, 57(2), 129–191.
- Kennedy, C., & McNally, L. (2005). Scale structure, degree modification, and the semantics of gradable predicates. *Language*, 345–381.
- Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, 71(2), 109–147.
- Schouwstra, M, Kirby, S & Culbertson, J. (2017). Silent gesture and noun phrase universals. In *Proceedings of the 39th Annual Conference of the Cognitive Science Society*. pp. 3095–3100.
- Toledo, A. & Sassoon, G. W. (2011). Absolute vs. relative adjectives: variance within versus between individuals. *Proceedings of Semantics and Linguistic Theory (SALT) 21*, 135–154.
- van Nipsen, K., van de Sandt-Koenderman, W. M. E., & Krahmer, E. (2017). Production and Comprehension of Pantomimes Used to Depict Objects. *Frontiers in Psychology*, 8, 1095. doi: 10.3389/fpsyg.2017.01095