Harmony in a non-harmonic language: word order learning in French children

Guillaume Braquet (s1577900@sms.ed.ac.uk)
Department of Linguistics and English Language, 3 Charles Street
Edinburgh, EH8 9AD UK

Jennifer Culbertson (jennifer.culbertson@ed.ac.uk)
Centre for Language Evolution, 3 Charles Street
Edinburgh, EH8 9AD UK

Abstract
Recent studies using artificial language learning have argued that the cross-linguistic frequency of harmonic word order patterns—in which heads are ordered consistently before or after dependents across syntactic categories—reflects a cognitive bias (Culbertson, Smolensky, & Legendre, 2012; Culbertson & Newport, 2015a). These studies suggest that English-speaking adults and children favor harmonic orders of nouns and different nominal modifiers (adjectives, numerals). However, because they target English learners, whose native language is harmonic in the nominal domain (Num-Adj-N), this preference may be based on transfer rather than a universal bias for harmony. We present new evidence from French-speaking children, whose native language is non-harmonic in this domain (Num-N-Adj). Our results reveal clear effects of native language transfer, but also evidence that a harmonic pattern is favored even in this population of learners.

Keywords: cognitive biases; artificial language learning; typology; syntax; word order; French

Introduction
Cross-category harmony (Greenberg, 1963; Hawkins, 1983) is perhaps the most well-known typological generalization of syntax: while it has been revised many times over the years (Dryer, 1992; Biberauer, Holmberg, & Roberts, 2014), most linguists continue to assume that it reflects some underlying “law of human behavior” (Greenberg, 1966). Nevertheless, like most typological generalizations, harmony is a statistical tendency, and researchers have pointed to the possibility that correlations in word order across categories may largely reflect the effects of language contact and shared inheritance rather than universal properties of human cognition (Dunn, Greenhill, Levinson, & Gray, 2011; Ladd, Roberts, & Dediu, 2014).

To provide evidence of a link to cognition, a number of recent studies have investigated word order harmony using artificial language learning experiments. Culbertson et al. (2012) taught English-speaking adults a miniature artificial language in which simple noun phrases, including a noun and either an adjective or a numeral word, are used to describe a set of pictures. The languages feature variable patterns of nominal word order that tend toward either harmonic or non-harmonic. Learners tended to regularize harmonic input patterns—using the input harmonic order more frequently than it appeared in the input. By contrast, they tended to shift non-harmonic patterns toward harmonic ones rather than regularizing. Culbertson and Newport (2015a) found an even stronger effect in English-speaking children, who almost across the board altered non-harmonic input patterns to make them harmonic (e.g., by changing the order of one of the modifiers to match the other). Indeed, even when children were taught a completely regular non-harmonic pattern (e.g., adjectives always follow and numerals always precede the noun), they still produced a harmonic output (Culbertson & Newport, 2015b).

Cognitive bias or native language influence?
These studies suggest the possibility that learners may preferentially change non-harmonic patterns to harmonic ones, rather than the reverse, offering a potential explanation for a similar typological asymmetry. Frequencies of the relevant nominal word order patterns are reported in the World Atlas of Language Structures Online, shown in Table 1. Harmonic patterns (both modifiers either pre-nominal or post-nominal) outnumber non-harmonic patterns. Note in addition, that within both harmonic and non-harmonic pattern types, the one with post-nominal adjectives is more frequent (i.e., N-Adj with N-Num; N-Adj with Num-N).

Table 1: Frequency of harmonic and non-harmonic combinations of noun with adjective and numeral (Dryer, 2013a, 2013b).

<table>
<thead>
<tr>
<th></th>
<th>Adj-N</th>
<th>N-Adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>Num-N</td>
<td>251</td>
<td>168</td>
</tr>
<tr>
<td>N-Num</td>
<td>37</td>
<td>509</td>
</tr>
</tbody>
</table>

Data from English-speaking learners reflects both a harmony bias and a bias for post-nominal adjectives to some extent. For example, English-speaking adults were equally likely to regularize pre- and the post-nominal harmonic patterns, despite the latter having less surface-level overlap with English than either non-harmonic pattern. They were also least likely to regularize the non-harmonic pattern with pre-nominal adjectives (Culbertson et al., 2012). English-speaking children did not exhibit any differences among non-harmonic patterns, however, their overall preference for harmonic patterns was in fact skewed toward the post-nominal N-Adj with N-Num (Culbertson & Newport, 2015a).

Nevertheless, evidence of a harmony preference in English speakers does not unequivocally suggest an explanation for the typological asymmetry seen here. This is of course because English itself exemplifies one of the two harmonic
patterns–Adj-N with Num-N. Thus a preference for harmony in this population may be due to transfer at a more abstract level. In other words, English speakers may find the post-nominal harmonic pattern easy to learn because they are used to treating numerals and adjectives similarly in terms of their relative order with the noun. If abstract transfer is responsible for the harmony bias in English speakers, then French speaking learners should behave quite differently. The default nominal word order in French exemplifies the more common non-harmonic pattern, N-Adj with Num-N. Numeral order is strictly pre-nominal, while adjectives are typically post-nominal but exhibit some flexibility. Most importantly, there is a small lexically-specified set of adjectives which precede the noun. Examples of default order are shown in (1).1

(1) a. maison bleue
house blue
‘blue house’

b. deux maisons
two houses
‘two houses’

French speakers could therefore be reasonably expected to prefer non-harmonic patterns over harmonic ones, since their prior language experience provides evidence that these two types of modifiers behave differently with respect to order. On the other hand, if the preference for harmony reflects a universal cognitive bias, then even French speakers–whose native language violates it–may exhibit its effects.

Experiment 1

Here we explore the potential effects of native language influence and cognitive biases on nominal word order learning in French-speaking children. We follow the general design and procedure of Culbertson and Newport (2015a) and Culbertson and Newport (2015b). Children are taught a variable version of one of four patterns corresponding to those in Table 1. They are trained on simple phrases comprising a noun with an adjective or a noun with a numeral, and are then tested on their production of those phrases. The extent to which learners accurately reproduce and regularize these variable patterns is used to infer their relative preferences.

Participants

Participants were 48 children (24 females), 6-7 years of age (mean=6;7, matched with participants in Culbertson and Newport (2015a)). They were recruited from elementary schools in Southwest France, and were native speakers of French, who were either monolingual, or bilingual in French and Occitan (a Romance language spoken in this region, which uses the same nominal word order as French). Parental consent was obtained for all participants. Three additional children were excluded from the analysis due to failure to complete the experimental session (2), or extremely low score on vocabulary learning (1).

Materials

The artificial language consisted of 10 words: 4 nouns corresponding to novel objects, 3 adjectives, and 3 numeral words. Following Culbertson and Newport (2015a), nouns were nonce words which were phonotactically plausible in the participants’ native language, and modifiers were pseudonone words resembling the corresponding native language words. This lexicon was used to describe pictures like those shown in Figure 1. Importantly, all of the adjectival modifiers used here appear in the default post-nominal order in French, they are not in the set of pre-nominally ordered adjectives.

Table 2: Artificial language lexicon.

<table>
<thead>
<tr>
<th>Nouns</th>
<th>Adjective</th>
<th>Numerals</th>
</tr>
</thead>
<tbody>
<tr>
<td>[bogi]</td>
<td>[bly] (bleu ‘blue’)</td>
<td>[doks] (deux, ‘two’)</td>
</tr>
<tr>
<td>[sefi]</td>
<td>[tachu] (tacheté, ‘spotted’)</td>
<td>[tura] (trois, ‘three’)</td>
</tr>
<tr>
<td>[voli]</td>
<td>[pølu] (poilu, ‘furry’)</td>
<td>[kitr] (quatre, ‘four’)</td>
</tr>
<tr>
<td>[kani]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1: Example visual stimuli.

Design & Procedure

Participants were randomly assigned to one of four input word order conditions. Each condition featured a dominant pattern for each modifier type, used in 75% of utterances. The remaining 25% used the alternative order. The variation present in the input was unpredictable; it was not conditioned on any particular lexical items. These conditions are illustrated in Table 3.

Table 3: Conditions, according to dominant pattern type (shaded cells are non-dominant).

<table>
<thead>
<tr>
<th>Harmonic dominant</th>
<th>Adj-N</th>
<th>N-Adj</th>
<th>Num-N</th>
<th>N-Num</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-N</td>
<td>75%</td>
<td>25%</td>
<td>75%</td>
<td>25%</td>
</tr>
<tr>
<td>Post-N</td>
<td>25%</td>
<td>75%</td>
<td>25%</td>
<td>75%</td>
</tr>
<tr>
<td>Non-harmonic dominant</td>
<td>N-Adj, Num-N</td>
<td>25%</td>
<td>75%</td>
<td>75%</td>
</tr>
<tr>
<td>Adj-N, N-Num</td>
<td>75%</td>
<td>25%</td>
<td>25%</td>
<td>75%</td>
</tr>
</tbody>
</table>

Following Culbertson and Newport (2015b), the experiment consisted of a single session, lasting approximately 30 minutes. During this session participants were trained and...
tested on the language with an experimenter present. The experiment was presented using PsychoPy software (Peirce, 2009) on a Macintosh laptop in a quiet corner of the child’s classroom. Children were told they would be playing a game to learn an alien language with the help of a friendly alien named Clémy.

The experiment began with a series of games designed to teach children the novel nouns and their meanings. The first game (20 trials) involved seeing a grayscale picture of a single object, listening to the label provided by the alien speaker, and repeating the label aloud. The second game (20 trials) involved listening to a label and clicking on whichever of the four novel objects it corresponded to. Feedback highlighting the correct picture was given on all trials. A sound indicated whether the child’s choice was correct or incorrect. The third game (20 trials) tested children’s ability to provide the correct label for an object shown on the screen. Feedback in the form of the correct noun label was provided on all trials.

The second part of the experiment consisted of a series of similar games designed to teach children simple phrases in the language. Recall that each phrase consisted of either a noun with an adjective or a noun with a numeral, but never both. Half of all trials featured an adjective, and half a numeral. The first game (48 trials) involved seeing pictures as in Figure 1 above, and hearing a phrase to describe it provided by the alien speaker. Participants were told to repeat the phrase. The second game (48 trials) involved hearing a phrase and choosing the corresponding picture from an array of four choices. Feedback was given as described above. Finally, children were shown a picture, and were asked to provide a description for it (48 trials). No feedback was given, however if the child had trouble with one of the words, the experimenter would help.

Results
Here we report results from the final phrase production task. Comprehension of phrases was uniformly high across conditions (> 85%). Figure 2 shows the average proportion of trials in which children used the dominant word order in each of the four word order input conditions. Children used the dominant order almost exclusively in the post-nominal harmonic condition, however dominant order use in the other three conditions was much lower. Children roughly matched the input proportions in the non-harmonic N-Adj, Num-N condition, which is most similar to their native language pattern.

Figure 3 shows use of the dominant pattern broken down by the type of modifier, illustrating more clearly what learners are producing when they don’t match the input. While in the post-nominal harmonic condition both modifier types are regularized equally often, there is a clear difference between modifier types for the other three conditions. In the pre-nominal harmonic condition, children tended to use the numeral in the dominant input order more often, however neither modifier type reproduces the input pattern closely. In the two non-harmonic conditions, children were more likely to match the dominant input order for whichever modifier tends to appear post-nominally: the adjective in the French-like N-Adj, Num-N condition, and the numeral in the Adj-N, N-Num condition. In other words, learners’ productions in these three conditions tended to move the language away from the input pattern and toward a post-nominal harmonic pattern.

These data were analyzed using mixed-effects logistic regression as implemented in the lme4 R package (Bates, 2010), with condition (input pattern type) and modifier type (adjective or numeral) as fixed effects and participants and items (stimulus picture) as random effects. Modifier type was sum coded; condition was treatment coded with the most French-like pattern (N-Adj, Num-N) as the baseline.
level. This model revealed a significant difference between the French-like (N-Adj, Num-N) condition and each of the other three conditions: compared to this condition there was significantly less use of the input order in the pre-nominal harmonic condition ($\beta = -3.99 \pm 0.56, p < 0.001$) and the non-harmonic Adj-N, Num-N condition ($\beta = -3.60 \pm 0.54, p < 0.001$), but significantly more use of the input order in the post-nominal harmonic condition ($\beta = 17.88 \pm 17.03, p = 0.03$). A significant main effect of modifier type was also present, indicating less matching of numeral order overall ($\beta = -2.28 \pm 0.22, p < 0.001$). Finally, significant interactions between condition and modifier type were found. As suggested by Figure 3, the main difference between the French-like and post-nominal harmonic conditions is in use of the dominant order for the numeral ($\beta = -12.50 \pm 17.04, p = 0.04$). By contrast, the opposite was true for both the pre-nominal harmonic ($\beta = 4.45 \pm 0.33, p < 0.001$) and non-harmonic Adj-N, Num-N ($\beta = 4.89 \pm 0.31, p < 0.001$) conditions where the bigger difference with the French-like condition was in use of the dominant adjective order.²

A visualization of the general direction of change from the input in each condition can be seen in Figure 4. This shows the proportion of pre-nominal adjectives and numerals each individual child produced, colored by their input condition. Learners generally cluster in the harmonic post-nominal area of the space, no learners for whom this was the input condition shifted away from this pattern. A smaller cluster appears around the non-harmonic French-like (N-Adj, Num-N) area, with children in that input condition plus some from the pre-nominal harmonic input condition who have switched only the adjective order to post-nominal. Figure 5 summarizes this picture by calculating each child’s preferred pattern, determined by the order used in the majority of utterances (greater than 50%) for each modifier type. For example, one child in the French-like N-Adj, Num-N condition produced adjectives post-nominally 100% of the time, and numerals post-nominally 88% of the time. This child was thus classified as having the post-nominal harmonic pattern.

Discussion

French-speaking children learning variable patterns of harmonic or non-harmonic nominal word order showed a striking pattern of behavior in their productions. When their input featured the post-nominal harmonic pattern N-Adj, Num-N as the dominant order, they regularized this pattern, producing nearly deterministic output. When their input featured a dominant pattern similar to their own native language, N-Adj, Num-N, they matched or regularized the adjective order, but not the numeral order. To the extent that children were fail-

²Note that the standard error on estimates for coefficients encoding contrasts between the French-like condition and the post-nominal harmonic condition is very high. This is likely due to individual variation in the French-like condition, where some participants strongly regularized the input order (similar to behavior in the pre-nominal harmonic condition), while others shifted away from the input dramatically. This is illustrated in Figure 4.
of the children in this experiment produced a pattern corresponding most closely to the post-nominal harmonic N-Adj, N-Num.

Recall that above we suggested two competing hypotheses which generated different predictions about the behavior of French-speaking learners in our experiment. Under the first hypothesis, influence from the native language is the primary driver of behavior. Previous results with English-speaking learners were consistent with abstract level transfer (i.e., a general preference for harmonic patterns). Thus French speakers would be predicted to prefer non-harmonic patterns, though a preference for the specific native language pattern (here N-Adj, Num-N) is also possible. The second hypothesis is that a cognitive bias favoring harmonic patterns is present universally across learners, even learners whose native language actively violates it. This predicts that French-speaking learners will, like English speakers, prefer harmonic patterns. The results reported here are not fully consistent with either hypothesis. Most obviously, learners did not show a preference for the pattern, or pattern type, most similar to their native language. Neither did they prefer harmonic patterns across the board. However, a combination of native language transfer effects and a universal harmony bias provides a coherent explanation. French-speaking children in our experiment exhibited a strong preference for post-nominal adjective order, in accord with their native language. This preference was generalized to numerals under a pressure for harmony.

In order for this explanation to work, we need to rule out a potential alternative: that French-speaking children age 6-7 years do not distinguish numerals from adjectives, or have not fully mastered the order of numerals in their native language. This could lead to the over-use of post-nominal ordering if children have, by contrast, mastered the default adjective order. There is relatively little work on the acquisition of nominal word order in monolingual French children. Indeed, we are not aware of corresponding work on the acquisition of numeral-noun order. We therefore extracted all instances of noun phrases including one of the numeral words ‘two’ through ‘ten’ in the Lyon corpus (Demuth & Tremblay, 2008). This is a publicly available corpus of naturalistic parent-child interactions including 5 children, recorded for 1 hour every 1-2 weeks from age 1 to 3 years (185 hours of speech total). Children’s first noun phrases with a numeral word occurred as early as 1;9. Out of a total of 258 instances, no word order errors were found. Based on this evidence, it appears likely that by 6-7 years of age, French-speaking children have long since mastered the pre-nominal order of numerals words in their language. Given the variation in adjective ordering in French, it would seem plausible that children may take longer to acquire this aspect of the syntax of their native language. However, evidence from spontaneous speech suggests that French-English bilingual children have mastered French adjective order very early as well, by 2;5 at least (Nicoladis, 2002). This may be accomplished earlier in monolinguals, in line with the early documented acquisition of adjective order in English by the age of 2 (Brown, 1973).

In our study then, it does not seem likely that French-learning children’s preference for post-nominal order in both adjectives and numerals is the result of a lack of knowledge about the syntax of these categories in their native language.

What remains then, is to understand how the results of this study fit with the previous findings from English-speaking adults and children. As mentioned above, English-speaking child learners did not generalize their native pre-nominal order preferentially, but rather readily produce both harmonic patterns. If anything, there was somewhat stronger preference for the post-nominal harmonic pattern (Culbertson & Newport, 2015a, 2015b). For English-speaking adults, both harmonic patterns were regularized, and among non-harmonic patterns the one with pre-nominal adjectives (Adj-N, N-Num), was particularly dispreferred (Culbertson et al., 2012). Recall that the typological distribution in Table 1 suggests both a preference for harmonic patterns, and a general preference for post-nominal adjectives. Thus among the harmonic patterns, the post-nominal one is more common, and within the non-harmonic patterns, N-Adj, Num-N is the most common. If these two pressures are at work across both learner populations, English and French, then we expect to see behavior mirror the typology. However, if these two biases are, as suggested above, influenced by learners’ native language experience, then a more complex picture emerges. For English-speaking children (and adults), the harmony bias is strengthened by native language experience, while the preference for post-nominal adjectives is weakened. As a result, learners strongly prefer harmonic languages, and only weakly prefer post-nominal over pre-nominal order. For French-speaking children, the opposite holds. The harmony bias is weakened, while the preference for post-nominal adjectives is strengthened. This results in a strong preference for the post-nominal harmonic pattern only, and changes to the input in each other condition which move the language toward that pattern via the adjective. Based on differences found between English-speaking adults and children, we would predict that French adults should show a similar but less dramatic pattern of preferences to children in our study.

Why these biases?

Both of the biases we have argued to be at work here reflect potentially quite general cognitive mechanisms. Harmonic patterns are simpler, in the sense that they involve fewer, more general rules, which can be generalized across categories. For extensive discussion of this idea, see Culbertson and Kirby (2016). The preference for post-nominal adjectives may reflect a pressure to establish the object of modification first, particularly in cases where the meaning of the noun is predictive in determining the meaning of the adjective, e.g., for gradable adjectives like ‘tall’ (Kamp & Partee, 1995; Ramscar, Yarlett, Dye, Denny, & Thorpe, 2010). Interestingly,
Nicoladis (2006) shows that French-English bilingual children (2;11-5;3) produce more reversals of adjective order than their monolingual counterparts in both French and English. These reversal errors were more likely to involve adjectives incorrectly in the post-nominal position than the reverse. In French, this corresponded to placing a typically pre-nominal adjective like grand ‘big’ after the noun. In English, this corresponded to placing adjectives in an incorrect post-nominal position when they were post-nominal in French (e.g., ‘a monkey purple’). This is in line with our findings in the sense that post-nominal adjective order seems to be more readily generalized than pre-nominal order. Overall then, the preference for post-nominal adjective ordering found among languages of the world and reflected in both English and French learners, may be related to general properties of learning and processing.

Conclusion

This study tested a classic hypothesis in linguistics: that harmonic word order patterns, which maintain a consistent order of syntactic heads relative to modifiers, are preferred to non-harmonic alternatives. Previous studies have shown that English-speaking learners prefer harmony in the nominal domain, however this could reflect abstract transfer since English exemplifies a pre-nominal harmonic pattern. Here, we targeted French-speaking child learners, whose native language is non-harmonic. If a harmony bias is present in this population it would provide a strong indication that this is indeed a universal pressure. By contrast, if transfer is the main driver of learning behavior, then no harmony bias is expected in this population—instead, non-harmonic patterns may be preferred. Our results revealed a strong preference for post-nominal harmonic order, which we argue reflects the effects of a harmony bias in conjunction with a preference for post-nominal adjectives.

Acknowledgments

We thank the schools, parents and children who participated in this study and the developmental linguistics reading group at the University of Edinburgh for comments on this work.

References


