The Recursive Set-Subset Principle orders adjectives and relative clauses in visual context

Adina Camelia Bleotu, Deborah Foucault & Tom Roeper*

Abstract We present evidence from two picture-matching experiments that native adult speakers of American English order and interpret sequences of nouns modified by adjectives and relative clauses containing adjectives in line with the RECURRENT SET-SUBSET ORDERING PRINCIPLE proposed in Bleotu & Roeper (2022a, b) and Bleotu, Foucault & Roeper (2023). This principle ensures an automatic mapping of set-subset semantics to a recursive syntax in Universal Grammar, such that set modifiers are merged closer to the noun than subset modifiers: leaves (N) that are short (SET) that are long (SUBSET) or short (SET) leaves (N) that are long (SUBSET). We here expand upon the interpretation and ordering of set-subset modification to elucidate the syntactic, semantic and pragmatic interface.

Keywords. English L1; adjectives; relative clauses; recursion; adjective ordering; modification; visual context; pragmatics

1. Introduction. Building upon previous findings (Bleotu & Roeper 2021a,b, Bleotu & Roeper 2022a, b, Foucault et al. 2022, Bleotu, Foucault & Roeper 2023), we embrace the claim that there is a perfect mapping between hierarchical syntax and recursive semantics, such that, in a sequence of two adjectives, the adjective merged closer to the noun maps onto a set interpretation, and the adjective merged further away from the noun maps onto a subset interpretation:

(1)  [long\text{subset} [short\text{set} leaves]]

We refer to this constraint on mapping as the RECURRENT SET-SUBSET ORDERING PRINCIPLE (RSSO), a principle which derives from the properties of Merge in Universal Grammar and provides an automatic set-subset ([subset [set]]) interpretation to the syntactic hierarchy:

(2)  RECURRENT SET-SUBSET ORDERING PRINCIPLE:
Set modifiers are merged earlier to the noun than subset modifiers.

Mapping a set-subset semantics onto a hierarchical syntax creates a recursive structure (Roepner 2011, Hollebrandse & Roeper 2014), which ultimately shows the human capacity to embed a category within the same category (Chomsky 1995; Hauser, Chomsky & Fitch 2002).

In the current paper, we ask whether, in a visual context involving pictures representing sets and subsets, adult native speakers of American English apply the RSSO in the syntax-semantic mapping to sequences of two relative clauses (RCs) containing Adj, as well as to a sequence of an Adj and an RC containing an Adj. Moreover, we ask if speakers associate these sequences with the corresponding recursive pictures.

(3)  a. [[[leaves [that are short]_{set}] [that are long]_{subset}]  
b. [[[short]_{set} leaves] that are long]_{subset}

* We would like to thank the participants of the study. In addition, we would also like to give thanks to the audiences at the Language Acquisition Research Center Seminar at UMass and the 2023 Annual Meeting of the LSA for their valuable comments and suggestions. Authors: Adina Camelia Bleotu, University of Bucharest (cameliableotu@gmail.com), Deborah Foucault, UMass Amherst (dfoucaulteth@umass.edu), Tom Roeper, UMass Amherst (roeper@linguist.umass.edu).

© 2023 Author(s). Published by the LSA with permission of the author(s) under a CC BY license.
Importantly, in Bleotu, Foucault & Roeper (2023), we showed that the RSSO applies to Adj/RC modifiers for purely linguistic stimuli, i.e., speakers tend to order and interpret RC and Adj modifiers as referring to the set if they are closer to the noun and as referring to the subset if they are further away from the noun. In contrast, we here combine linguistic stimuli with pictures, probing into whether adult native speakers of American English consider the set-subset ordering when mapping the set-subset order/interpretation onto visual stimuli, i.e., actual pictures of leaves that are short that are long.

2. **Background on the RSSO.** Much of the research on adjectives has focused on adjective ordering restrictions (AORs), which have been investigated in multiple theoretical and experimental studies (e.g., Dixon 1982; Matthei 1982; Cinque 1994, 2005, 2010; Scott 2002; Bryant 2006; Scontras et al. 2017, a.o.). Many of the discussions around AORs address the question of whether speakers order adjectives depending on the cognitive dimensions they express and their degree of subjectivity (Scontras et al. 2017, 2019), and/or whether they do it on the basis of an innate syntax. The general consensus reached by the literature on this topic is that, at least in American English, speakers seem to observe an ordering such as the one in (4).

(4) QUALITY > SIZE > SHAPE > COLOR > PROVENANCE
(Sproat & Shih 1991; Dixon 1982; Scott 2002)

This ordering seems to be correlated with the subjectivity expressed by the adjective: adjectives expressing more objective properties tend to be placed closer to the noun, and adjectives expressing more subjective properties tend to be placed further away from the noun (Scontras et al. 2017, 2019). It is unclear whether this (cognitive) tendency is hardwired syntactically or not, since there is a lot of crosslinguistic variation in this domain. AORs seem to be quite strict in English but more flexible in languages such as Greek (Leivada & Westergaard 2019), Hebrew (Trainin & Shetreet 2021) or Romanian (Cornilescu & Cosma 2019; Cornilescu & Giurgea, 2013; Cornilescu & Nicolae 2016; Luciu & Bleotu to appear; Truşcă & Bleotu to appear). Moreover, even English seems to be rather flexible in how it orders shape and color adjectives (see Grohe & Schulz 2021).

While much of the literature on the ordering of adjectives has focused on how participants order adjectives in the absence of a set-subset context, we focus instead on how adult participants order adjectives in the presence of such a context. Experimental research by Bleotu & Roeper (2021a, b) has shown that Romanian adult native speakers are able to produce and interpret recursive adjectives according to a set-subset distinction. Romanian is a noun-adjective language, which exhibits a mirror order of English. In a context involving referring to subsets within sets, adults understand the adjective closest to the noun as referring to the set and the subsequent adjective as referring to the subset. Thus, when familiarized with multiple big and small flowers, as in Figure 1, Romanian adult native speakers are able to understand that a sequence such as (5) refers to the subset of small flowers from the set of big flowers (represented by the groups labelled 3 and 4 in Figure 1), while a sequence such as (6) refers to the sub-subset big flowers from the subset of small flowers from the set of small flowers (represented by the group labelled 7 in Figure 1).

![Figure 1. Picture employed in Bleotu & Roeper (2021a, b)](image)
(5) 2-Level Adj Recursion:
flori mari mici (3&4)
flowers big small
‘small big flowers’

(6) 3-Level Adj Recursion:
flori mici mici mari (7)
flowers small small big
‘big small small flowers’

Foucault et al. (2022) showed that American English-speaking adults and (73%) 4-7-year-old children are able to produce and interpret sequences such as *big small mushrooms, small big mushrooms*. For instance, *small big mushrooms* is understood as referring to the subset of small mushrooms from the set of big mushrooms.

Bleotu & Roeper (2022 a, b) recently showed that Romanian-speaking adults and 4-year-old children observe the RSSO in contexts which align with cognitive adjective orders: they associate a picture depicting a subset of long leaves from a set of green leaves with the sequence long green leaves rather than green long leaves. Interestingly, they also seem to observe the RSSO in contexts which go against cognitive adjective orders: they associate a picture depicting a subset of green leaves from a set of long leaves with the sequence green long leaves rather than long green leaves. This shows that the RSSO is more fundamental to the grammar than cognitive adjective restrictions.

We advocate that the RSSO maps the set-subset semantics onto recursive syntax by first merging set adjectives with the noun and then subset adjectives. This can be implemented in several ways. One such way is to treat adjectives as specifiers of special functional projections dedicated to set and subset properties (see 7), in the cartographic approach, à la Cinque (1994, 2005, 2010). Another way is to adopt an adjunction account (Abels & Neeleman 2010; Kremers 2003), and successively merge set and subset adjectives to the left as modifiers of the noun (see 7).

(7) a. *Cartographic Approach*  
   
   ![Cartographic Approach Diagram](image)

   On the basis of two multiple choice linguistic tasks, Bleotu, Foucault & Roeper (2023) recently showed that the RSSO also applies to RCs containing adjectives and to combinations of adjectives and RCs, such as (8).

(8) a. [[leaves [that are short]<sub>set</sub> [that are long]<sub>subset</sub>]
   b1. [[short]<sub>set</sub> leaves] that are long<sub>subset</sub>
   b2. [short<sub>subset</sub> [leaves that are long<sub>set</sub>]]
Native speakers of American English merge Adjs and RC set modifiers closer to the noun and Adj and RC subset modifiers further away from the noun. Moreover, they interpret modifiers closer to the noun as denoting sets and modifiers further away from the noun as denoting subsets. Interestingly, the RSSO clearly predicts that in a N RC RC sequence, the RC merged closest to the N should denote the set and the RC further away should denote the subset, which is consistent with the findings in Bleotu, Foucault & Roeper (2023). However, for A N RC sequences, the RSSO is compatible both with an interpretation where the Adj is the set and the RC is the subset (8b1), as well as with an interpretation where the RC is the set and the Adj is the subset (8b2), given that either the Adj or the RC could merge with the N. For this structure, Bleotu, Foucault & Roeper (2023) find that participants interpret the Adj as the set and the RC as the subset (as on 8b1).

3. Experiment. We extend our investigation of the RSSO by means of two picture-matching tasks testing how native speakers of American English order RCs and Adjs, and whether they observe RSSO for RCs, just as for Adjs, and interpret the closest modifier as a set. Importantly, while Bleotu, Foucault & Roeper (2023) investigated this matter on the basis of purely linguistic items, we approach it from the perspective of mapping a picture onto a linguistic sequence.

3.1. Participants. A test group of 19 adult native speakers of American English took part in the task.

3.2. Predictions. In terms of ordering, we make different predictions depending on whether the RSSO holds universally or not.

We first consider N RC RC sequences. If RSSO is a Universal Grammar constraint on ordering modifiers, then it should hold for RCs as well, not just for Adjs. We thus expect RCs to have the inverse order of prenominal adjectives based on Adj being prenominal and RCs being postnominal in English. If, however, RCs are ordered via symmetric coordination or derived via a freer form of extraposition, then participants might order RCs differently than Adjs. For instance, if a participant moves the RC referring to the set to the left of the RC referring to the subset, then the structure in (9) becomes (10):

(9)  [[leaves [that are short]_set] [that are long]_subset]
(10) Move RC_set to the left of RC_subset → [[leaves [that are long]_subset] [that are short]_set]

For A N RC sequences, both the Adj and the RC are close to the noun, so, in principle, the RSSO might give rise either to the A_set N RC_subset order or the A_subset N RC_set. However, if other considerations come into play, such as the compound-like nature of Adj N formations or the lightness of the Adj compared to the heaviness of the RC, then we might expect the order A_set N RC_subset, as in Bleotu, Foucault & Roeper (2023).

In terms of interpretation, we expect participants to associate modifiers closer to N with sets. For N RC RC sequences, this means that the first RC will be interpreted as referring to a set, whereas the second RC will be interpreted as referring to a subset. For A N RC, this means that, in principle, either the A or the RC could be interpreted as referring to the set. However, as already discussed, if other considerations come into play, then we might expect A N RC to be interpreted as A_set N RC_subset, as in Bleotu, Foucault & Roeper (2023).

1 Importantly, while A N RC sequences are in principle ambiguous between two interpretations, this is not the case for N RC sequences, where the N is a compound consisting of an Adj and a noun. For a sequence such as blueberries that are big, the Adj is an integral part of the noun and is not subject to a compositional interpretation. In such a sequence, the RC that are big is a Set modifier of the N.
Given that the task involved picture-matching, we might expect additional complications in ordering and interpreting nominal sequences modified by Adjs or RCs containing Adjs. The challenge of ordering and interpretation is now not merely a linguistic one, as in Bleotu, Foucault & Roeper (2023), but it also involves associating pictures with sequences of words. Thus, in addition to the syntax-semantics mapping guided by the RSSO, participants also must consider an independent representation of context and pragmatically map expressions to context or context to expressions.

3.3. **PROCEDURE.** Participants responded to two successive forced choice picture-matching tasks:

(i) **an Ordering Task** (see Table 1) consisting of 16 questions in which participants had to choose which expression has the same meaning as a picture representing an A₂ A₁ N sequence (A₁ N RC₂ or A₂ N RC₁; N RC₁ RC₂ or N RC₂ RC₁); and

(ii) **an Interpretation Task** (see Table 2) consisting of 24 questions in which participants had to choose the picture representing the set from which the character should pick certain items (e.g., can long leaves that are short be found among long leaves or short leaves)

We organized the experiment such that all Ordering Task questions precede the Interpretation Task to avoid a possible priming effect of interpretation on ordering.

Importantly, in the Ordering Task participants must map expressions to pictures, versus in the Interpretation Task, they must map pictures to expressions. The direction of the linguistic-visual context mapping thus varies between tasks. Participants can rely on linguistic contrast between two expressions in the Ordering Task, while they can rely on visual contrast between two pictures in the Interpretation Task.

In terms of categories, in the Ordering Task we tested A₂ A₁ N and A₁ N RC₂ (Adjective + Noun + Relative Clause) equivalences and A₂ A₁ N and N RC₁ RC₂ (Noun + Relative Clause + Relative Clause) equivalences. In the Interpretation Task, we tested three combinations containing Adjs (A₂ A₁ N), RCs (N RC₁ RC₂), and Adjs/RCs combined (A₁ N RC₂).

In terms of the cognitive dimension expressed by the modifiers, we tested both structures containing:

(i) Size-Size combinations (short-long, small-big, short-tall, thin-fat) and

(ii) Size-Color combinations (long-green, big-red, thin-brown, tall-yellow).

<table>
<thead>
<tr>
<th>Ordering Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tom likes the circled leaves, see? There are many leaves: a pile of long leaves (the first four leaves) and a pile of short leaves (the last four leaves). Are the circled leaves…?</td>
</tr>
</tbody>
</table>

| A N RC | i) short leaves that are long  
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ii) long leaves that are short</td>
</tr>
</tbody>
</table>
| N RC RC| i) leaves that are long that are short  
|        | ii) leaves that are short that are long |

Table 1. Examples of experimental items with size-size modifiers from the Ordering Task
3.4. RESULTS. Regarding ordering, participants mostly chose orders which represent mirrors of \([long_{subset} [short_{set} \text{ leaves}]]\). In particular, they chose \([[short_{set} \text{ leaves}] \text{ that are } long_{subset}]] \text{ of the time and } [[\text{ leaves that are short}_{set} \text{ that are long}_{subset}]] 68.4\% \text{ of the time.}

In accordance with RSSO, they also mostly interpreted the modifier closest to the noun as defining the \textbf{set}, and the subsequent modifiers branching further away from the N defining the \textbf{subset}. In particular, for the sequence \([long_{subset} [short_{set} \text{ leaves}]]\) they interpreted the Adj closest to the noun (A\(_1\)) as defining the \textbf{set} 80\% of the time. For the sequence \([[\text{ leaves that are short}_{set} \text{ that are long}_{subset}]]\) they interpreted the RC closest to the noun (RC\(_1\)) as set-defining 73.68\% of the time. For the sequence \([[short_{set} \text{ leaves} \text{ that are long}_{subset}]]\), they interpreted the Adj as defining the \textbf{set} 76.31\% of the time.

An ANOVA test reveals no significant differences in ordering or interpretation between A N RC sequences and N RC RC sequences. Moreover, Size-Size modifiers and Size-Color modifiers seem to pattern alike.

Interestingly, when taking a look at the individual data we see a lot of intra-individual variation. For each of the combinations of modifiers, participants seem to also allow for a certain proportion of orders or interpretations which are not necessarily in line with the RSSO. Please note, however, that, in the case of A N RC, any order/interpretation is basically in line with the RSSO.

4. Discussion. Our results show that the RSSO seems to cut across categories, ordering both Adjs and RCs in line with set-subset considerations. Nevertheless, the variation in the individual data suggests that participants’ preferences are not so firm. This could potentially be explained by participants experiencing additional challenges in ordering/interpreting sequences of nouns and recursive modifiers when pictures come into play. Such an explanation is supported by looking comparatively at the results from the current experiment and the results from Bleotu, Foucault & Roeper (2023): we find that, when participants have to make decisions of ordering/interpreting recursive modifiers exclusively on the basis of linguistic items, they seem to provide more answers that are in line with the RSSO (> 70\% in the Ordering Task, > 90\% in the Interpretation Task). Thus, the presence of pictures that participants must match to recursive sequences seems to create complications in mapping the set-subset semantics onto a hierarchical
In effect, to account for the impact of the visual representations we have to assume that participants have a representation of the context in which the set-subset relations are expressed. In the Ordering Task, participants are exposed to a picture and can build on linguistic contrast to select the adequate expression to refer to it. In the Interpretation Task, participants are exposed to one expression and can build on visual contrast to select the right picture corresponding to it. The contrast between two expressions associated with a picture or two pictures associated with one expression may influence participants to choose the unusual non-recursive alternative, which is possible though not frequent ([short\textsubscript{subset} [leaves that are long\textsubscript{set}], [[leaves [that are long]\textsubscript{subset} [that are short]]\textsubscript{set}]).

Although participants’ answers are not at ceiling, the RSSO seems to guide the ordering of recursive modifiers overall. When a noun is modified by two Adjs or two RCs, the modifier closest to the noun is interpreted as the set and the modifier further away from the noun is interpreted as the subset. However, when a noun is modified by an Adj and an RC, Adjs seem to merge to the noun first, while RCs merge onto the Adj+N sequence. Thus, although either the Adj or the RC could in principle be associated with a set interpretation, Adjs seem to have priority. This could be because, unlike RCs, adjectives can merge directly to bare nouns and create compound-like formations. It could also be because Adjs are lighter than RCs, and a preference to place heavier constituents at the end of a sequence may come into play (Arnold et al. 2000). However, please note that, with the adequate context and/or intonation, we can also get an interpretation where the RC is the set and the Adj is the subset. This suggests that, at a deeper level, an A N RC sequence may actually be ambiguous between A\textsubscript{set} N RC\textsubscript{subset} and A\textsubscript{subset} N RC\textsubscript{set}, but, once the choice between the two is made on the basis on the context and/or intonation, the RSSO makes a clear prediction with respect to interpretation: the closest modifier is the set, and the one further away is the subset.

In what follows, we discuss two possible theoretical implementations of our findings: an adjunction parameter (branching directionality) account (Kremers 2003; Abels & Neeleman 2010) and a movement account (Cinque 1994, 1995, 2010).

According to the adjunction parameter (branching directionality) account, Adjs and RCs have different parameter settings: while Adjs are adjoined to the left, RCs are adjoined to the right, as shown in (11).

\begin{equation}
\begin{array}{ccc}
\text{N RC\textsubscript{set} RC \textsubscript{subset}} & \text{N RC \textsubscript{subset} N RC \textsubscript{set}} & \text{N RC \textsubscript{subset} N RC \textsubscript{set}} \\
\text{A\textsubscript{set} N} & \text{A\textsubscript{set} N} & \text{A\textsubscript{set} N} \\
\text{NP} & \text{NP} & \text{NP} \\
\text{AP\textsubscript{subset} long} & \text{NP} & \text{NP} \\
\text{AP\textsubscript{subset} short} & \text{NP} & \text{NP} \\
\text{N leaves} & \text{RC\textsubscript{subset} that are long} & \text{RC\textsubscript{subset} that are long} \\
\end{array}
\end{equation}

According to the movement account, the orders we find are derived from the basic order FP\textsubscript{subset} FP\textsubscript{set} – as shown in (12). In order to obtain the order [[A\textsubscript{set} N] RC\textsubscript{subset}], we start from RC\textsubscript{subset} A\textsubscript{set} N and move [A\textsubscript{set}+N] to SpecFP\textsubscript{subset}. In order to obtain the order [[N RC\textsubscript{set}] RC\textsubscript{subset} set]], we start from RC\textsubscript{subset} RC\textsubscript{set} N and do the operation of Roll-Up: we first move N to SpecFP\textsubscript{set} and we then move [N+RC\textsubscript{set}] to SpecFP\textsubscript{subset}. 

\begin{equation}
\begin{array}{ccc}
\text{NP} & \text{NP} & \text{NP} \\
\text{AP\textsubscript{subset} long} & \text{NP} & \text{NP} \\
\text{AP\textsubscript{subset} short} & \text{NP} & \text{NP} \\
\text{N leaves} & \text{RC\textsubscript{subset} that are long} & \text{RC\textsubscript{subset} that are long} \\
\end{array}
\end{equation}
Both accounts are able to capture the fact that RCs seem to show a mirror order of Adjs in American English. Moreover, both accounts can capture the existence of adjectival mirror effects in other languages such as Italian or Spanish: either by assuming adjunction or Roll-Up movement. However, an adjunction parameter account in terms of branching directionality (Kremers 2003; Abels & Neeleman 2010) is structurally simpler: one can set the parameter for branching directionality differently across languages and within the same language.\(^2\)

Since a grammar which employs less movement is preferable to one employing more movement, we argue that an adjunction parameter account can better capture our findings.

5. Conclusion. In the current paper, we have presented experimental evidence from two picture-matching tasks in favor of the existence of a recursive set-subset ordering constraint which orders adjectives (Adjs) and relative clauses (RCs). This constraint is a fundamental UG principle which maps the set-subset semantics onto syntax through the operation of Merge: set modifiers are merged first to the noun, followed by subset adjectives. While our participants do show a preference for observing this constraint, the existence of a significant number of answers that are not in line with the RSSO may be explained via challenges in the mapping of linguistic expressions and the representation of pictures. Thus, participants have to map structures, meanings, and visual context onto each other, which ultimately shows that they all share a set-subset representation. In future work, we plan to extend this line of research to American English children, investigating whether they order and interpret recursive Adjs/ RCs in line with the RSSO.

References


Bleotu, Adina Camelia & Tom Roeper. 2021a. Small big flowers or small and big flowers? Simple is better and Roll-Up is too complex for Romanian 5-year-olds. In Danielle Dionne &

\(^2\) Please note, however, that the opposite argument could also be made, namely that the Roll-Up movement account presents the advantage of not requiring different parameter settings within a grammar, but of imposing a consistent rule in UG.


Bleotu, Adina Camelia, Deborah Foucault & Tom Roeper. 2023. *Long short leaves* are ‘leaves that are short that are long’. The Recursive Set-Subset Principle orders both adjectives and relative clauses. To appear in *Chicago Linguistic Society (CLS)*.


