Phrase Structure Basics

Ad Neeleman (UCL) (with Klaus Abels, Peter Ackema, Zoë Belk, Vikki Janke, Misako Tanaka, Amanda Payne, Joy Philip and Hans van de Koot)

1. Linearization

While linearization is variable, certain movements must be leftward. We argue that this is because rightward movement of obligatory material causes parsing difficulties that are otherwise absent. Evidence comes from heavy-XP shift. The resulting constraint can be brought to bear on the linear asymmetry described in Greenberg's (1963) Universal 20 if difficulties in parsing drastically suppress the typological frequency (as per Hawkins 1990, 2009 and Kirby 1999).

This section is based on Klaus Abels and Ad Neeleman. 2024. *e.* Ms. UCL.

1.1 Nature abhors a vacuum

Neeleman and Weerman (1999) argued for two basic syntactic freedoms: freedom of linearization (as in Dutch PP extraposition) and freedom of projection (as in Dutch scrambling). The approach is that flexibility comes for free, while rigidity requires an explanation.

The analysis of Universal 20 requires both freedom of linearization and rigid leftward movement of (constituents containing) the noun (Cinque 2005, Abels and Neeleman 2012):

(1) When any or all of the items demonstrative, numeral, and descriptive adjective precede the noun, they are always found in that order. If they follow, the order is either the same or its exact opposite. (Greenberg 1963)

(2) *Symmetric base-generation:*
   a. [Dem [Num [A N]]] attested
   b. [[[N A] Num] Dem] attested

(3) *Asymmetric movement:*
   a. [N [Dem [Num [A ~~N~~]]]] attested
   b. *[[[[~~N~~ A] Num] Dem] N] unattested

We show that certain rightward movements are problematic in parsing and propose that a subset of structures derived by such movements are subject to typological attrition if given enough time.

If traces are (partial) copies of moved categories, the parser cannot insert a trace before having hypothesized that a given category has undergone movement. This implies filler-driven parsing of leftward movement (Phillips and Wagers 2007).

On the filler-driven strategy, incremental parsing of the string in (4) may be conceptualized as in (5) (the symbol ***e*** represents an empty position). The thing to note that ***e*** is filled immediately after it has been created (i.e. before D is parsed).

(4) X-A-B-C-t_X-D-E-F

(5) a. X Leftward movement
   b. X A B C
   c. X A B C ***e***
   d. X A B C ~~X~~
   e. X A B C X D E F

Contrast this with the parsing steps involved in incremental analysis of a string in which an element X that would normally obligatorily follow C has moved rightward:

(6) A-B-C-t_X-D-E-F-X

(7) a. A B C — Rightward movement (obligatory material)
b. A B C ***e***
c. A B C ***e*** D E F
d. A B C ***e*** D E F X
e. A B C ~~X~~ D E F X

The sequence of steps in (7) can in turn be distinguished from those necessary to parse the same string if X is an element that can *optionally* follow C.

(8) a. A B C — Rightward movement (optional material)
b. A B C D E F
c. A B C D E F X
d. A B C ***e*** D E F X
e. A B C ~~X~~ D E F X

Thus, rightward movement of obligatory material is qualitatively different from both leftward movement and rightward movement of optional material. This qualitative difference creates difficulties in the parsing of rightward movement of obligatory material.

The core of the problem is that in (7) ***e*** fails to be filled and is removed from the leading edge of the parse tree. There is no problem with empty positions that remain unfilled as long as they are pushed rightward as further material is processed:

(9) a. A B C ***e***
b. A B C D ***e***
c. A B C D E ***e***
d. A B C D E F

(10) Leading Edge Constraint (LEC)
e must be filled immediately or remain at the leading edge of the parse tree.

The LEC may explain why the movement component in the analysis of Universal 20 is asymmetric. The noun is an obligatory element in the noun phrase. Hence, the movement in (3b) will require a parsing process in which ***e*** is generated and subsequently pushed from the leading edge. If so, language change away from the offending grammar is likely (Hawkins 2009 and Kirby 1999).

How to test the proposal? Heavy-XP shift is relevant. If the verb is obligatorily transitive, heavy XP shift should cause parsing difficulties:

(11) a. John ate (the food that his brother prepared). — optionally transitive
b. John ate t_1 yesterday [the food that his brother prepared]$_1$.

(12) a. John devoured *(the food that his brother prepared). — obligatorily transitive
b. John devoured t_1 yesterday [the food that his brother prepared]$_1$.

(13) a. John ate — Heavy XP shift (optional object)
b. John ate yesterday
c. John ate yesterday the food that his brother prepared
d. John ate ***e*** yesterday the food that his brother prepared
e. John ate ~~the food that his brother prepared~~ yesterday the food that his brother prepared

(14) a. John devoured — Heavy XP shift (obligatory object)
b. John devoured ***e***
c. John devoured ***e*** yesterday
d. John devoured ***e*** yesterday the food that his brother prepared
e. John devoured ~~the food that his brother prepared~~ yesterday the food that his brother prepared

(15) Heavy-XP shift of obligatory objects should cause parsing difficulties during the processing of material crossed by the movement; heavy-XP shift of optional objects should not.

There is a further prediction. After the filler has been identified, there is one subsequent step in (13), namely the placement of a copy in ***e***. In (14) there are *two* subsequent steps, as ***e*** must first be created before it can be filled by a copy of the shifted object. Thus:

(16) The processing of an obligatory object that has undergone heavy-XP shift should be easier than the processing of an optional object that has undergone this movement.

We consider these predictions before returning to Universal 20.

2. Testing the predictions

2.1 Literature review

Two sources of convergent evidence: corpus studies by Wasow (1997a,b) and eye-tracking experiments by Staub et al. (2006). Corpus studies are relevant because speakers parse their own speech (Levelt 1989). Wasow found the following (V*t* = obligatorily transitive verb; V*p* = transitive verbs that have an alternative use with an immediately following PP and no NP object.

Brown corpus (written) $\chi^2(1) = 6.49, p < .02$		Switchboard corpus (spoken) $\chi^2(1) = 10.65, p < .01$	
V*t*	V*p*	V*t*	V*p*
5.6%	9.3%	1.45%	3.82%

Table 1: Percentage of heavy-XP-shifted DP-objects in two corpora as determined by verb transitivity

Two eye-tracking experiments reported in by Staub et al. 2006 confirm the predictions in (15) and (15) more directly:

	Verb	Adverbial	Object
V_{OblTr}	no difference	slower on all measures	faster on crucial measures
V_{OptTr}		faster on all measures	slower on crucial measures

Table 2: Comparison of reading times in three regions of sentences with a heavy-XP-shifted object

The effect is categorical rather than gradient. The verbs that Staub et al. used in their second experiment showed variation in their preference for a transitive frame. However, the reported effect was not correlated with this, but was only found with obligatorily transitive verbs

Given the importance of predictions (15) and (16) for our argumentation, we ran two additional experiments, intended to replicate and extend Staub et al.'s findings using different experimental paradigms (self-paced reading and a maze task):

	one adverbial	two adverbials
V_{OblTr}	Subject V_{OblTr} *t* Adv_{prim} Object	Subject V_{OblTr} *t* Adv_{sec} Adv_{prim} Object
V_{OptTr}	Subject V_{OptTr} *t* Adv_{prim} Object	Subject V_{OptTr} *t* Adv_{sec} Adv_{prim} Object

Table 3: Structures tested in the self-paced reading and maze tasks

The point of the second adverbial was to determine whether any slow-down on the material separating verb and object could be due to surprisal. The Leading Edge Constraint predicts that parsing difficulties should persist. By constrast, if the slow-down reported by Staub et al. is an effect of surprisal, reaction times should return to normal as the point of surprisal recedes.
Thus, the predictions under scrutiny are the following:

(17) a. Reaction times over the primary adverbial will be slower in the obligatorily transitive condition than in the optionally transitive condition, irrespective of the presence of absence of a delay adverbial.
b. Reaction times over the first few words of the shifted object will be faster in the obligatorily transitive condition than in the optionally transitive condition.

2.2 Self-paced reading

In self-paced reading (Just et al. 1982) participants read a sentence on a computer screen. The sentence starts off masked and the participant presses a button to reveal each successive word and mask the

previous word. The time between button presses is measured.

We ran our experiment online using Gorilla and Prolific.

Materials consisted of ninety-six test sentences and forty-eight fillers. Test sentences were identical to or adapted from those used by Staub et al. (2006). A sample item made up of four test sentences is given in (18) (where *attack* is optionally and *bother* obligatorily transitive). Fillers were grammatical sentences (comparable in complexity) in which some category had been fronted or extraposed.

(18) a. Sara attacked$_1$ with no mercy$_2$ the red ants$_4$ living on the windowsill.
b. Sara attacked$_1$ for ten minutes$_3$ with no mercy$_2$ the red ants$_4$ living on the windowsill.
c. Sara bothered$_1$ with no mercy$_2$ the red ants$_4$ living on the windowsill.
d. Sara bothered$_1$ for ten minutes$_3$ with no mercy$_2$ the red ants$_4$ living on the windowsill.

Thus, the experiment had a 2x2 design, with the nature of the verb (obligatorily/optionally transtive) and the presence/ absence of an delay adverbial as the variables we manipulated and reading times over the primary adverbial and the onset of the heavy NP as dependent variables.

- Prediction (17a) is confirmed if (i) reading times over the primary adverbial are slower in the obligatorily transitive conditions and (ii) there is no interaction with the presence of a delay adverbial.
- Prediction (17b) is confirmed if reading times at the onset of the heavy NP are faster in the obligatorily transitive conditions.

Test items were distributed over four list. In addition, all fillers were added to all lists. Fillers and test sentences were presented in pseudo-randomized order. All experimental and filler sentences were followed by a comprehension question.

We recuited 116 participants. For various reasons, thirty-four participants were excluded, leaving eighty-two participants whose data was subjected to analysis. Data cleaning yielded 101,695 response times, which were log-transformed and residualized. Filler trials were stripped out and the resulting 16,961 responses were subjected to analysis.

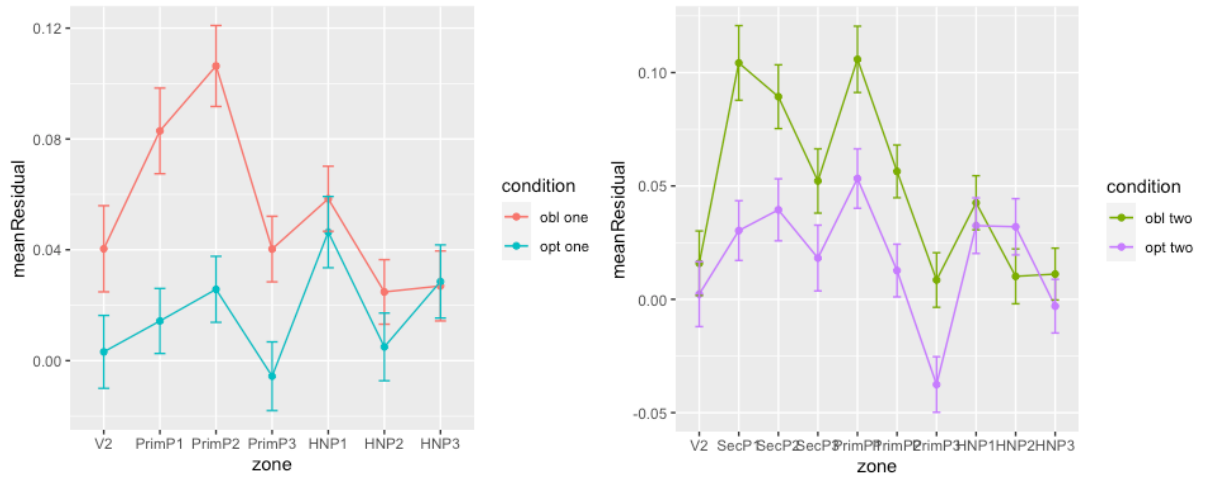


Table 4: Plots for reading times in the one/two adverbial conditions (*obl*: obligatorily transitive verb; *opt*: optionally transitive verb; *one*: one intervening adverbial; *two*: two intervening adverbials)

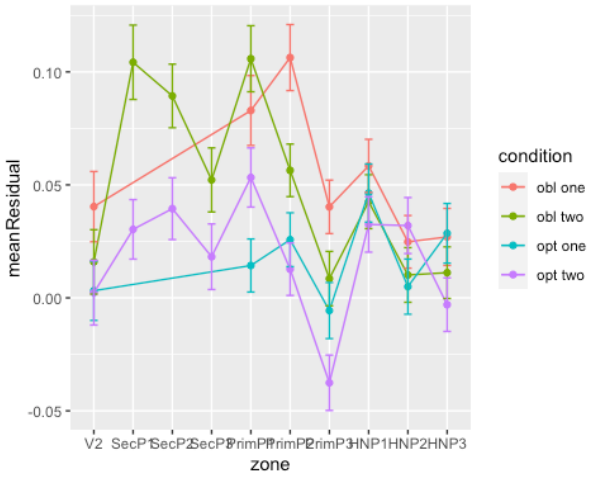


Table 5: Combined reading time plots (*obl*: obligatorily transitive verb; *opt*: optionally transitive verb; *one*: one intervening adverbial; *two*: two intervening adverbials)

As the graphs show, residual reading times are longer for obligatorily transitive verbs across both the primary and the delay adverbial, with spikes at the onset of each adverbial.

We fitted seven linear mixed effects models to the residual reading times with contrast-coded conditions for the number of adverbials and the optional/obligatory transitivity of the verb.

Significance testing was done using model reductions. A limited random effects structure was used to ensure that all models, including reduced ones, would converge non-defectively. We used Bonferroni correction in determining significance levels (as seven models were fitted).

```
## # A tibble: 28 × 8
##    Zone   term           estimate std.error statistic     df p.value
corrected.p
##    <fct>  <chr>             <dbl>     <dbl>     <dbl>  <dbl>   <dbl> <chr>
##  1 V2     (Intercept)      0.0159    0.0149    1.07     50.2 0.291   ns
##  2 V2     OBL_contrast     0.0255    0.0136    1.87   1890.  0.0622  ns
##  3 V2     PPS_contrast    -0.0126    0.0137   -0.925  1891.  0.355   ns
##  4 V2     OBL_contrast:…  -0.0214    0.0273   -0.784  1889.  0.433   ns
##  5 PrimP1 (Intercept)      0.0652    0.0127    5.12     31.7 0.00001 ***
##  6 PrimP1 OBL_contrast     0.06      0.0135    4.46   1896.  0.00001 ****
##  7 PrimP1 PPS_contrast     0.0302    0.0135    2.24   1897.  0.0252  ns
##  8 PrimP1 OBL_contrast:…  -0.0158    0.0269   -0.586  1895.  0.558   ns
##  9 PrimP2 (Intercept)      0.0506    0.0097    5.19     30.3 0.00001 ****
## 10 PrimP2 OBL_contrast     0.0621    0.0123    5.04   1895.  0       ****
## 11 PrimP2 PPS_contrast    -0.0311    0.0123   -2.52   1897.  0.0117  ns
## 12 PrimP2 OBL_contrast:…  -0.036     0.0246   -1.46   1894.  0.144   ns
## 13 PrimP3 (Intercept)      0.0015    0.0102    0.152    32.2 0.881   ns
## 14 PrimP3 OBL_contrast     0.0457    0.0119    3.84   1896.  0.00013 ***
## 15 PrimP3 PPS_contrast    -0.0319    0.0119   -2.68   1898.  0.00747 ns
## 16 PrimP3 OBL_contrast:…   0         0.0238    0.0011 1895.  0.999   ns
## 17 HNP1   (Intercept)      0.0451    0.0099    4.54     30.0 0.00009 ***
## 18 HNP1   OBL_contrast     0.0108    0.012     0.896  1898.  0.370   ns
## 19 HNP1   PPS_contrast    -0.0148    0.012    -1.23   1898.  0.219   ns
## 20 HNP1   OBL_contrast:…  -0.0029    0.024    -0.120  1897.  0.905   ns
## 21 HNP2   (Intercept)      0.0182    0.0122    1.49     35.5 0.145   ns
## 22 HNP2   OBL_contrast    -0.0013    0.0117   -0.112  1892.  0.911   ns
## 23 HNP2   PPS_contrast     0.0064    0.0117    0.55   1892.  0.582   ns
## 24 HNP2   OBL_contrast:…  -0.0425    0.0234   -1.82   1891.  0.0691  ns
## 25 HNP3   (Intercept)      0.0174    0.0151    1.16     30.8 0.256   ns
## 26 HNP3   OBL_contrast     0.0059    0.0118    0.498  1895.  0.618   ns
## 27 HNP3   PPS_contrast    -0.0217    0.0118   -1.85   1896.  0.0652  ns
## 28 HNP3   OBL_contrast:…   0.0145    0.0236    0.614  1895.  0.539   ns
```

Table 6: Significance levels of reactions times in verb, the primary adverbial and the heavy object

Our results show a significant main effect of the optional/obligatory transitivity of the verb on the primary adverbial (PrimP1, PrimP2 and PrimP3), but no main effect of the presence of a delay adverbial and no interaction. So, the slowdown over the primary adverbial in the transitive conditions occurs independently of whether this adverbial immediately follows the verb or appears further downstream. No other effects were significant: obligatory transitivity did not affect reading times over the onset of the heavy NP. Hence, the data confirm prediction (17a), but not prediction (17b).

An additional analysis that included the delay adverbial (ten models, the three for the delay adverbial missing all one-adverbial conditions) found a significant effect of obligatory transitivity on the three

words of the primary adverbial. In addition, it found a significant slowdown in the obligatorily transitive condition on the first word of the delay PP, in line with expectations.

2.3 Maze

Our second experiment was a maze task (Forster et al. 2009). In a maze task participants are presented with a forced choice at each position between a target word that allows a legitimate continuation of the sentence and a distractor that does not. Participants press a button corresponding to what they think is the correct word, with reaction time used as the dependent measure.

Distractors were existing words chosen Boyce et al.'s (2020) A(uto)-Maze method and subsequently checked and adjusted by hand.

We ran the experiment online using Gorilla and Prolific. Target sentences were largely identical to those used in the self-paced reading study, but we increased the number of words preceding the verb as errors can occur in maze tasks at the beginning of sentences. In the sample in (19), *attack* is optionally and *bother* obligatorily transitive.

(19)

a.		b.		c.		d.	
The	x-x-x	The	x-x-x	The	x-x-x	The	x-x-x
woman	yours	woman	yours	woman	yours	woman	yours
in	mid	in	mid	in	mid	in	mid
the	than	the	than	the	than	the	than
old	nor	old	nor	old	nor	old	nor
house	rates	house	rates	house	rates	house	rates
attacked$_1$	patience	attacked$_1$	patience	bothered$_1$	patience	bothered$_1$	patience
with$_2$	blog	for$_3$	tool	with$_2$	blog	for$_3$	tool
no$_2$	ha	ten$_3$	app	no$_2$	ha	ten$_3$	app
mercy$_2$	tends	minutes$_3$	improve	mercy$_2$	tends	minutes$_3$	improve
the$_4$	cent	with$_2$	blog	the$_4$	cent	with$_2$	blog
red$_4$	ago	no$_2$	ha	red$_4$	ago	no$_2$	ha
ants$_4$	vivo	mercy$_2$	tends	ants$_4$	vivo	mercy$_2$	tends
living	videos	the$_4$	cent	living	videos	the$_4$	cent
on	gain	red$_4$	ago	on	gain	red$_4$	ago
the	glad	ants$_4$	vivo	the	glad	ants$_4$	vivo
windowsill.	stipulates	living	videos	windowsill.	stipulates	living	videos
		on	gain			on	gain
		the	glad			the	glad
		windowsill.	stipulates			windowsill.	stipulates

Fillers were grammatical sentences with a category fronted or extraposed and with a complexity comparable to that of the test items.

Thus, the maze experiment had a 2x2 design (verb: obligatorily or optionally transitive; delay adverbial: present or absent). We measured reaction times over the primary adverbial and onset of the heavy NP, which allowed us to test the predictions in (17).

The test items were distributed over four list in Latin-square fashion. In addition, the lists contained all forty-eight fillers. Fillers and test sentences were presented in pseudo-randomized order.

We recruited 119 participants, of which six were excluded. The data generated by the remaining 113 participants was submitted to further analysis.

Data were excluded for all sentences in which a participant responded incorrectly at any word. This left 15,017 data points. Table 7 shows separate plots for the conditions with one and two intervening adverbials, respectively. Table 8 shows the combined plot.

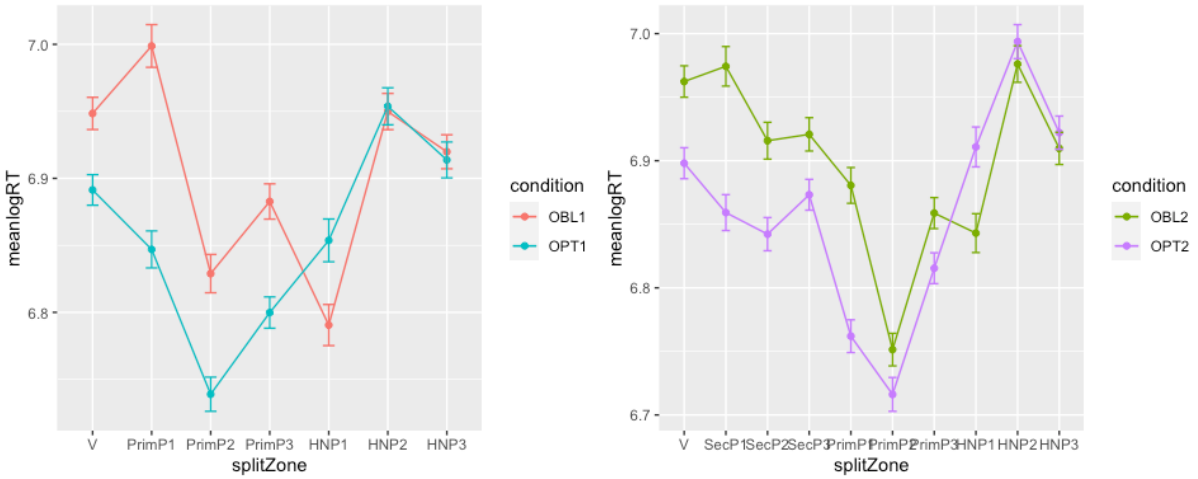


Table 7: Plots for choice times in the one/two adverbial conditions (*obl*: obligatorily transitive verb; *opt*: optionally transitive verb; *1*: one intervening adverbial; *2*: two intervening adverbials)

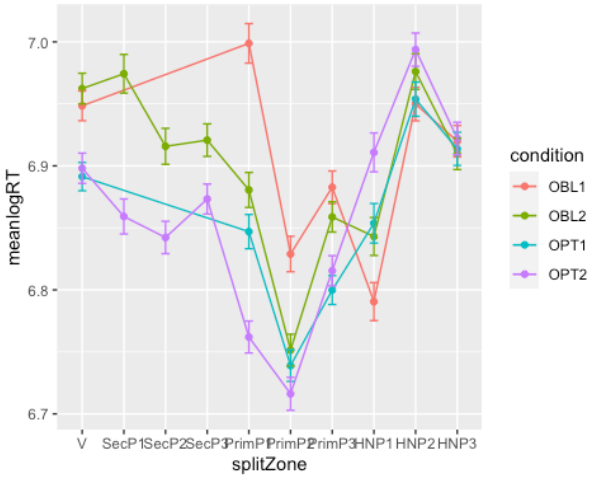


Table 8: Combined choice time plots (*obl*: obligatorily transitive verb; *opt*: optionally transitive verb; *1*: one intervening adverbial; *2*: two intervening adverbials)

Two aspects of the plots stand out. In both the one-adverbial and the two-adverbial condition the line representing items with obligatorily transitive verbs is systematically above the line representing items with optionally transitive verbs in the first area of interest (PrimP1–PrimP3). In the second area of interest, the lines cross on HNP1 and then effectively merge. In the the two-adverbial condition, the delay adverbial is also slower if the verb is obligatorily transitive.

For analysis, conditions were contrast coded. Log-transformed reaction times of these data were analyzed at each word (excluding the delay adverbial) in seven separate linear mixed effects models. Significance was tested using model reduction and is reported with Boneferroni correction (see Table 9).

Our findings fully bear out the predictsion in (16). We found a significant main effect of obligatory transitivity for the primary adverbial (a slowdown over PrimP1, PrimP2 and PrimP3) and a significant main effect of obligatory transitivity on the onset of the shifted NP (a speeding up overHNP1). We also found a significant main effect of the presence of a delay adverbial on PrimP1, PrimP2, HNP1, and HNP2, but crucially no significant interactions with obligatory transitivity. Thi lack of interactions suggests that the slowdown in the obligatorily transitive conditions should not be attributed to surprisal (but rather to the LEC).

The LEC also predicts that obligatory transitivity will affect reading times over the delay adverbial. We ran an additional analysis that included the delay adverbial, so that we could check this prediction. This analysis found a significant main effect of obligatory transitivity on SecP1, SecP2, PrimP1, PrimP2, and HNP1 and nowhere else. In addition, we found a significant main effect of presence of the delay adverbial on PrimP1, PrimP2, HNP1, and HNP2, but crucially no significant interaction with obligatory transitivity.

```
## # A tibble: 28 × 8
##    Zone   term          estimate std.error statistic     df  p.value
corrected.p
##    <fct>  <chr>            <dbl>     <dbl>     <dbl>  <dbl>    <dbl> <chr>
##  1 V      (Intercept)    6.93e+0    0.0230  301.       55.0 4.06e-90 ****
##  2 V      OBL_contrast  -7.56e-2    0.0296   -2.55     23.2 1.77e- 2 ns
##  3 V      PP_contrast    1.56e-2    0.0116    1.34   1654.  1.80e- 1 ns
##  4 V      OBL_contrast… -2.08e-2    0.0231   -0.898  1651.  3.69e- 1 ns
##  5 PrimP1 (Intercept)    6.89e+0    0.0263  262.       55.4 2.10e-87 ****
##  6 PrimP1 OBL_contrast  -1.51e-1    0.0229   -6.59     24.0 8.05e- 7 ****
##  7 PrimP1 PP_contrast   -1.07e-1    0.0144   -7.42   1629.  1.94e-13 ****
##  8 PrimP1 OBL_contrast…  5.54e-2    0.0287    1.93   1629.  5.40e- 2 ns
##  9 PrimP2 (Intercept)    6.77e+0    0.0408  166.       30.3 1.96e-46 ****
## 10 PrimP2 OBL_contrast  -6.51e-2    0.0152   -4.27     23.7 2.71e- 4 **
## 11 PrimP2 PP_contrast   -4.90e-2    0.0118   -4.15   1670.  3.52e- 5 ***
## 12 PrimP2 OBL_contrast…  6.09e-2    0.0235    2.59   1673.  9.66e- 3 ns
## 13 PrimP3 (Intercept)    6.84e+0    0.0259  264.       55.3 1.97e-87 ****
## 14 PrimP3 OBL_contrast  -6.79e-2    0.0230   -2.95     23.5 7.04e- 3 *
## 15 PrimP3 PP_contrast    2.11e-4    0.0118    0.0179 1661.  9.86e- 1 ns
## 16 PrimP3 OBL_contrast…  3.18e-2    0.0235    1.35   1661.  1.76e- 1 ns
## 17 HNP1   (Intercept)    6.86e+0    0.0313  219.       46.0 3.98e-71 ****
## 18 HNP1   OBL_contrast   7.98e-2    0.0226    3.53     24.0 1.72e- 3 *
## 19 HNP1   PP_contrast    5.74e-2    0.0153    3.75   1599.  1.80e- 4 **
## 20 HNP1   OBL_contrast… -2.60e-2    0.0305   -0.853  1601.  3.94e- 1 ns
## 21 HNP2   (Intercept)    6.99e+0    0.0339  206.       35.9 7.57e-57 ****
## 22 HNP2   OBL_contrast   1.71e-2    0.0177    0.965    22.2 3.45e- 1 ns
## 23 HNP2   PP_contrast    3.82e-2    0.0133    2.87   1600.  4.16e- 3 *
## 24 HNP2   OBL_contrast…  3.18e-2    0.0265    1.20   1603.  2.29e- 1 ns
## 25 HNP3   (Intercept)    6.93e+0    0.0300  231.       38.0 1.90e-61 ****
## 26 HNP3   OBL_contrast   1.82e-2    0.0181    1.01     22.7 3.24e- 1 ns
## 27 HNP3   PP_contrast   -2.21e-2    0.0130   -1.70   1617.  8.89e- 2 ns
## 28 HNP3   OBL_contrast…  3.61e-2    0.0259    1.40   1619.  1.63e- 1 ns
```

Table 9: Significance levels of reactions times in verb, the primary adverbial and the heavy object

We have contrasted LEC effects (which persists) with effects of surprisal (which should subside over time). Interestingly, the error rates contain a pattern suggestive of surprisal. Our error rate analysis consisted of 21,413 datapoints from all four zones of interest (674 datapoints were excluded). Plots are given in table 10 (for the one-adverbial and two-adverbial conditions separately) and table 11 (combined).

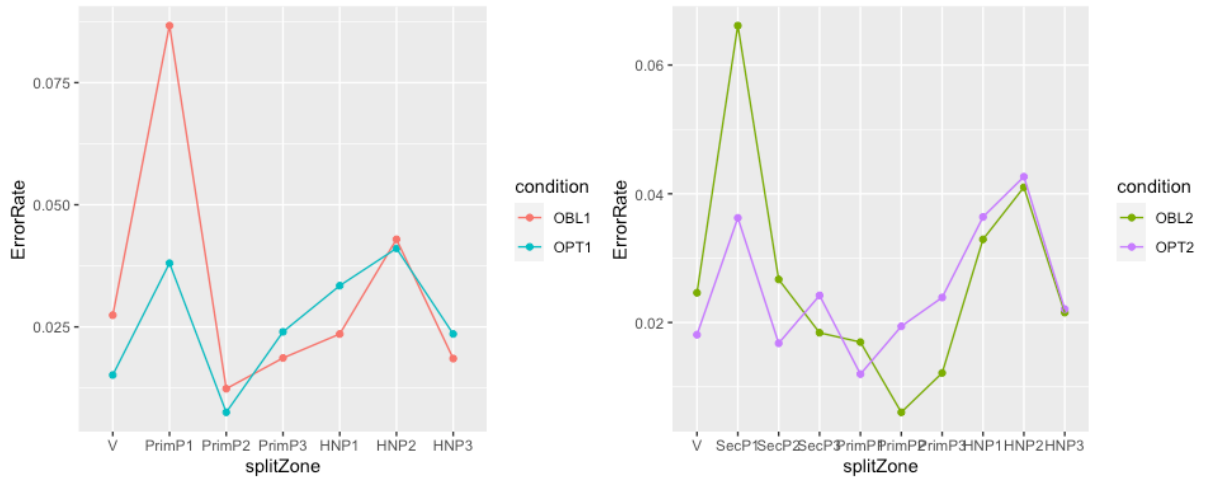


Table 10: Plots for error rates in the one/two adverbial conditions (*obl*: obligatorily transitive verb; *opt*: optionally transitive verb; *1*: one intervening adverbial; *2*: two intervening adverbials)

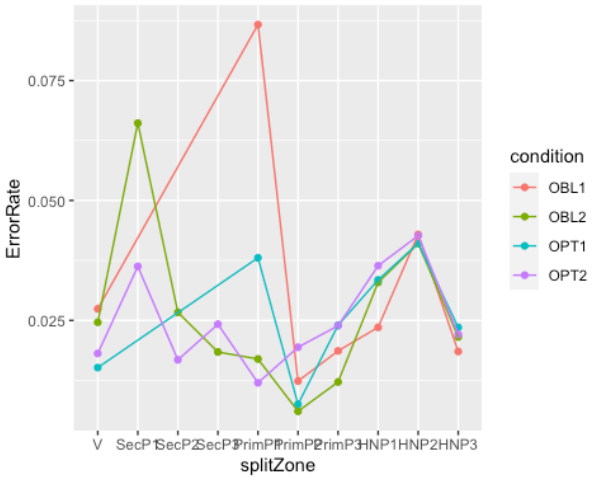


Table 11: Combined error rate plots (*obl*: obligatorily transitive verb; *opt*: optionally transitive verb; *1*: one intervening adverbial; *2*: two intervening adverbials)

We excluded data for the delay adverbial. Conditions of interest were contrast coded. Error rates were analyzed by conducting 7 separate logistic regressions. Significance was evaluated using Bonferroni correction. The results are summarized in table 12 below.

The only significant effect was a main effect of the number of adverbials on PrimP1, due to strongly elevated error levels at PrimP1 in the one-PP condition, independent of the obligatory or non-obligatory transitivity of the verb.

The lack of any spike in error rates over the primary adverbial in the two-adverbial condition suggests that the peak in error rates after the verb is a consequence of surprisal.

The peak in error rates was present whether the verb was obligatorily or optionally transitive. This again suggests that we are dealing with surprisal. In Staub et al.'s first experiment, the optionally transitive verbs had a high propensity for transitive use. As we based our experimental items on Staub et al.'s, we inherited this transitivity bias.

```
## # A tibble: 28 × 7
##    Zone   term                 estimate std.error statistic  p.value
corrected.p
##    <fct>  <chr>                   <dbl>     <dbl>     <dbl>    <dbl> <chr>
##  1 V      (Intercept)              4.81     0.366   13.1    2.14e-39 ****
##  2 V      OBL_contrast             0.465    0.290    1.60   1.09e- 1 ns
##  3 V      PP_contrast             -0.0399   0.288   -0.138  8.90e- 1 ns
##  4 V      OBL_contrast:PP_con…    -0.393    0.590   -0.667  5.05e- 1 ns
##  5 PrimP1 (Intercept)              4.12     0.290   14.2    8.44e-46 ****
##  6 PrimP1 OBL_contrast             0.666    0.273    2.44   1.47e- 2 ns
##  7 PrimP1 PP_contrast              1.55     0.274    5.65   1.62e- 8 ****
##  8 PrimP1 OBL_contrast:PP_con…    -0.586    0.543   -1.08   2.81e- 1 ns
##  9 PrimP2 (Intercept)              5.57     0.502   11.1    1.44e-28 ****
## 10 PrimP2 OBL_contrast            -0.352    0.411   -0.856  3.92e- 1 ns
## 11 PrimP2 PP_contrast             -0.128    0.411   -0.311  7.56e- 1 ns
## 12 PrimP2 OBL_contrast:PP_con…    -1.74     0.822   -2.12   3.44e- 2 ns
## 13 PrimP3 (Intercept)              4.30     0.264   16.3    1.10e-59 ****
## 14 PrimP3 OBL_contrast            -0.473    0.294   -1.61   1.08e- 1 ns
## 15 PrimP3 PP_contrast              0.220    0.294    0.750  4.54e- 1 ns
## 16 PrimP3 OBL_contrast:PP_con…    -0.429    0.587   -0.731  4.65e- 1 ns
## 17 HNP1   (Intercept)              4.39     0.313   14.0    1.35e-44 ****
## 18 HNP1   OBL_contrast            -0.240    0.246   -0.976  3.29e- 1 ns
## 19 HNP1   PP_contrast             -0.259    0.248   -1.04   2.96e- 1 ns
## 20 HNP1   OBL_contrast:PP_con…    0.374     0.492    0.760  4.47e- 1 ns
## 21 HNP2   (Intercept)              3.69     0.241   15.3    4.60e-53 ****
## 22 HNP2   OBL_contrast             0.0152   0.208    0.0734 9.41e- 1 ns
## 23 HNP2   PP_contrast             -0.0128   0.207   -0.0618 9.51e- 1 ns
## 24 HNP2   OBL_contrast:PP_con…    -0.0914   0.415   -0.220  8.26e- 1 ns
## 25 HNP3   (Intercept)              4.34     0.299   14.5    1.06e-47 ****
## 26 HNP3   OBL_contrast            -0.120    0.279   -0.432  6.66e- 1 ns
## 27 HNP3   PP_contrast             -0.0406   0.278   -0.146  8.84e- 1 ns
## 28 HNP3   OBL_contrast:PP_con…    0.235     0.555    0.423  6.73e- 1 ns
```

Table 12: Significance levels of error rates in the verb, the primary adverbial and the heavy object

The fact that error rates decrease over time, while the slowdown in reaction times persists suggests that the maze experiment allows us to tease apart the consequences of surprisal and the consequences of the LEC. But this implies that the latter cannot be reduced to the former.

3. Universal 20

3.1 The pattern

Universal 20 is a description of attested and non-attested neutral orders:

(20) A given order as neutral if it is (i) the only grammatical order, or (ii) the only grammatical order that can be used in an out-of-the-blue context, or (iii) the most frequent grammatical order that can be used in an out-of-the-blue context.

The complete paradigm of all attested and nonattested neutral orders of demonstrative, numeral, adjective and noun was described in Cinque 2005, 2014:

	I	II	III	IV
a.	Dem Num A N	N A Num Dem	N Dem Num A	A Num Dem N
b.	Dem Num N A	A N Num Dem	Dem N Num A	A Num N Dem
c.	Dem A N Num	Num N A Dem	A N Dem Num	Num Dem N A
d.	Dem N A Num	Num A N Dem	N Num A Dem	Dem A Num N
e.	A Dem Num N	N Num Dem A	N Dem A Num	Num A Dem N
f.	A Dem N Num	Num N Dem A	N A Dem Num	Num Dem A N

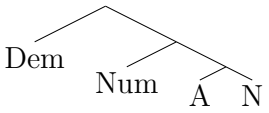
Table 13: Attested and unattested orders in the noun phrase according to Cinque 2005, 2013

Abels and Neeleman (2012) propose a reinterpretation of Cinque's (2005) antisymmetric analysis of the data in Table 13. First, the following universal hierarchy holds:

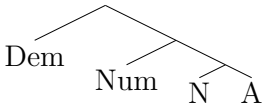
(21) Dem > Num > A > N

If we allow variation in linearization, this allows for the following representations:

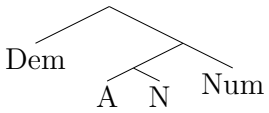
(22) Ia.



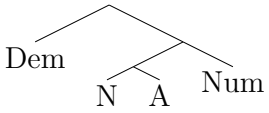
Ib.



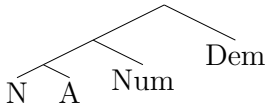
Ic.



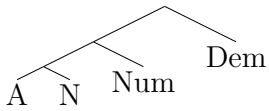
Id.



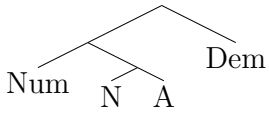
IIa.



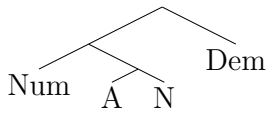
IIb.



IIc.

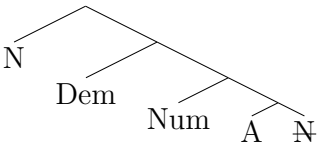


IId.

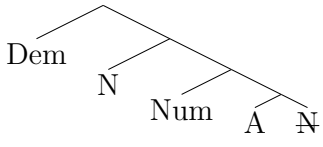


What needs to be added to capture the full paradigm is that neutral word orders can be derived by leftward movement of the noun or a constituent containing noun and adjective:

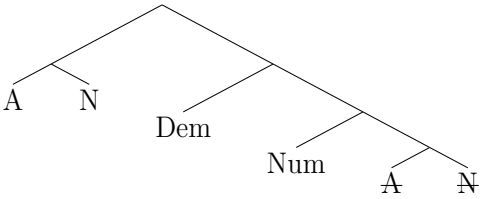
(23) IIIa.



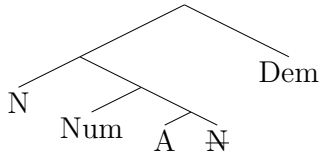
IIIb.



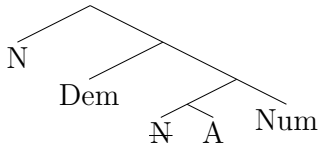
IIIc.



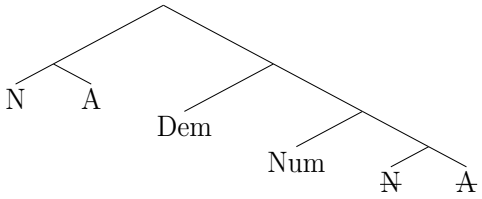
IIId.



IIIe.



IIIf.



We next consider why the unattested orders are underivable. Given (21), N and A must be adjacent in the base. Since the noun cannot move rightward, it is impossible to separate N and A if they come in N-A order (see (24a)). Similarly, given that numerals must be adjacent in the base to the substring comprising the N and A, (24b) is excluded. Hence (Ie,f) and (IVa-f) are ruled out.

(24) a. *A … X … N
b. *Num … X … A+N

Two unattested orders remain, namely (IIe,f). Here, the adjective and the noun are separated, suggesting that N has moved. But if so, the base structures for (IIe) and (IIf) must have been either *Num-Dem-A-N or *Num-Dem-N-A (compare (IVc,f)).

In conclusion, the constraint that movements deriving neutral order cannot be rightward is instrumental in explaining the full pattern of attested and nonattested orders in the extended nominal projection. The crucial question at this point is *why* movement of the noun must be leftward. Our suggestion is that this is due to parsing difficulties.

3.2 Synchronic and diachronic consequences of the Leading Edge Constraint

We have argued that rightward movement of obligatory categories creates parsing difficulties (because of the LEC. Evidence comes from heavy XP shift. This factor can explain the absence of rightward movement orders in the U20 domain:

- Difficulty in parsing corresponds to typological frequency. Hawkins 2004: "Grammars have conventionalized syntactic structures in proportion to their degree of preference in performance, as evidenced by patterns of selection in corpora and by ease of processing in psycholinguistic experiments".
- The mechanism for this has been proposed by Kirby 1999. Kirby distinguishes input and intake. The language-learning child is confronted with utterances in its environment (the input). She must parse those utterances to come to representations that can be used for the acquisition of grammar (the intake). Structures that are hard to parse will have a reduced frequency in the intake compared to their frequency in the input. Kirby shows that over time this will eliminate structures that are hard to parse (assuming variation).
- This mechanism straightforwardly militates against the relevant rightward movements in the extended nominal projection; they systematically cause parsing difficulties and will therefore be suppressed in the child's intake.
- The logic just outlined leads to an urgent question. Why should the very same parsing difficulties give rise to apparent ungrammaticality in the noun phrase and to no more than reduced frequency in the case of heavy-XP shift of obligatory objects?
- There is in fact an obvious distinguishing factor. Rightward movement in the extended nominal projection will always cause parsing difficulties. However, heavy-XP shift has instances that do not cause parsing difficulties, namely when the verb only optionally selects an object. This difference is crucial in acquisition. Heavy-XP shift of optional objects implies that there are relevant inputs in the intake that are – we propose – sufficiently frequent to acquire a general process of rightward shift (note that many transitive verbs are optionally transitive). This general process can subsequently apply to optional and obligatory objects alike.

3.3 The frequency of Universal 20 orders

The proposed account of Universal 20 based on the notion that the LEC reduces the typological frequency of structures derived by rightward movement of the noun to zero. This is a departure from the accounts in Cinque 2005 and Abels and Neeleman 2012, where the ban on rightward noun movement was taken to be grammatical and hence unviolable. How does the LEC fit in with other factors affecting frequency?

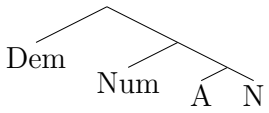
The typological distribution of Universal 20 orders follows a power law, with some orders much more

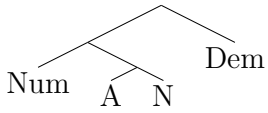
frequent than others. A good fit with the known distribution can be achieved if there is (i) a preference for harmonic orders, (ii) a preference for N-A order, and (iii) a preference for non-movement orders, evaluated in this sequence. Like other accounts of typological frequency based on formal factors, the above presupposes a certain degree of 'linguistic entropy'.

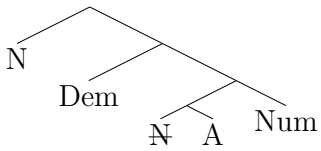
- The preference for nonmovement structures is tied to the widely accepted notion that movement is a costly operation.
- Our definition of harmony as it applies to extended projections is given in (25).

(25) *Harmony*
An extended projection is harmonic if nonprojecting categories are either uniformly located on right branches or uniformly located on left branches.

The definition in (25) characterizes the structure in (25a) as harmonic and the structure in (25b) as disharmonic. Any structure in which the noun moves is also disharmonic (this logic extends to structures in which the noun-adjective unit moves if, as Georgi and Müller's (2010) argue, it reprojects in its landing site). Harmony is presumably a preference in acquisition (Culbertson et al. 2012).

(26) a.


b.


c.


- The preference for N-A order holds of the base positions of these elements. Evidence for it comes from an artificial language learning experiment reported in Culbertson et al. 2012. Culbertson et al. attribute this to Kamp and Partee's (1995) (Semantic-)Head Primacy Principle, which states that the noun is interpreted first, followed by the adjective, regardless of syntactic word order. This is because the interpretation of (gradable) adjectives depends on the noun (a big butterfly is a butterfly that is big for a butterfly).

  Thus, if the noun comes first, the adjective can be interpreted immediately, but if the adjective comes first, it has to be buffered for interpretation until after the noun is interpreted, and so the latter order is dispreferred as less efficient in parsing.

We can understand the ranking of the fourteen Universal 20 orders in table 13 if we assume that following hierarchy of preferences holds:

(27) Harmony > N-before-A > No-Movement

The effects are given in table 14. (i) Harmonic orders outrank disharmonic orders. (ii) Within the class of harmonic orders N-A-Num-Dem is more frequent than Dem-Num-A-N, and within the class of disharmonic orders, those in which the adjective is not (or not necessarily) merged to the left of the noun are more frequent than those in which the adjective *is* unambiguously merged to the left of the noun. (iii) Within each of the two disharmonic classes, nonmovement orders outrank movement orders.

	Harmony	*N-before-A*	*No-Movement*	*Predicted ranking*
N-A-Num-Dem				1
Dem-Num-A-N		*		2
Dem-N-A-Num	*			3-5
Dem-Num-N-A	*			3-5
Num-N-A-Dem	*			3-5
N-A-Dem-Num	*		*	6-10
Dem-N-Num-A	*		*	6-10
N-Num-A-Dem	*		*	6-10
N-Dem-A-Num	*		*	6-10
N-Dem-Num-A	*		*	6-10
Dem-A-N-Num	*	*		11-13
Num-A-N-Dem	*	*		11-13
A-N-Num-Dem	*	*		11-13
A-N-Dem-Num	*	*	*	14

Table 14: Ranking of Universal 20 orders as predicted by (27)

Table 15 gives several measures of typological frequency in the Universal 20 domain. These do not always agree. Nonetheless, the predicted ranking is a good match with the rankings based on each of the individual measures (mismatches appear in bold), and a perfect match with the average rankings.

	AF(D)	*LF(D)*	*GF(D)*	*LF(C)*	*GF(C)*	*Spr*	*A.R.*	*P.R.*
N-A-Num-Dem	1 (44.17)	1 (182)	1 (85)	1 (630)	1 (136)	1	1 (1)	1
Dem-Num-A-N	2 (35.56)	2 (113)	2 (57)	2 (442)	2 (115)	1	2 (2)	2
Dem-N-A-Num	3 (29.95)	4 (53)	3 (40)	4 (204)	3 (89)	2	3 (3.4)	3-5
Dem-Num-N-A	4 (22.12)	5 (40)	4 (32)	5 (184)	4 (76)	2	4-5 (4.4)	3-5
Num-N-A-Dem	**6 (14.54)**	3 (67)	5 (27)	3 (239)	5 (49)	4	4-5 (4.4)	3-5
N-A-Dem-Num	**5 (14.08)**	6 (36)	6 (19)	6 (103)	6 (35)	2	6 (5.8)	6-10
Dem-N-Num-A	7 (9.75)	8 (12)	8 (10)	10 (50)	9 (29)	4	7 (8.4)	6-10
N-Num-A-Dem	8-9 (9.00)	10 (11)	9 (9)	8 (70)	8 (34)	3	8 (8.7)	6-10
N-Dem-A-Num	8-9 (9.00)	7 (13)	7 (11)	**12 (32)**	**12 (17)**	6	9-10 (9.3)	6-10
N-Dem-Num-A	10 (5.67)	**11-12 (8)**	**11 (6)**	7 (83)	7 (25)	6	9-10 (9.3)	6-10
Dem-A-N-Num	11 (5.34)	**9 (12)**	**10 (7)**	11 (48)	**10 (27)**	3	11 (10.2)	11-13
Num-A-N-Dem	12 (4.00)	11-12 (8)	12 (5)	**9 (55)**	11 (21)	4	12 (11.1)	11-13
A-N-Num-Dem	13 (3.00)	13-14 (5)	13-14 (3)	13 (33)	13 (13)	2	13 (13.2)	11-13
A-N-Dem-Num	14 (2.50)	13-14 (5)	13-14 (3)	14 (20)	14 (8)	2	14 (13.8)	14

AF(D): adjusted frequency (Dryer 2018); LF(D): language frequency (Dryer 2018); GF(D): genus frequency (Dryer 2018); language frequency (Cinque 2023); GF(C): genus frequency (Cinque 2023); Spr: spread of rankings; AR: average ranking over the five measures; PR: predicted ranking according to table 14.

Table 15: A comparison of the predicted ranking and the actual ranking based on five measures

Given that ranking of factors is necessary, we may assume that the Leading Edge Constraint is simply stronger than the other factors that reduce typological frequency:

(28) LEC > Harmony > N-before-A > No-Movement

It is not straightforward how to distinguish highly marked orders from ungrammatical orders at the tail end of the distribution (where data are sparse and misanalysis is more likely). Hence, we do not know whether any language has LEC violating orders. This is a matter for further research.

4. Conclusion

- Universal 20 pattern emerges from the interaction of two factors: variable linearization and leftward movement of (a constituent containing) the noun.
- The ban on rightward movement finds its explanation in parsing difficulties associated with rightward movement of obligatory material.

Appendix: A more detailed model of the parser

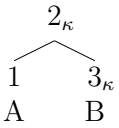
1. General

We develop a more detailed model of the parser and and discuss how it applies to heavy-XP shift. The output of the parser is a tree. The process of tree construction for expository purposes treated as a one-track process) is incremental and monotonic:

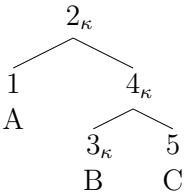
(29) a. *Incrementality*: The parser assigns each input symbol a structural position as soon as it comes in.
b. *Monotonicity*: No commitment present at stage *n* of the parsing process may be abandoned at stage *n*+1.

The input string is structured using three basic relations: domination, precedence and headedness. The notion of headedness we will use is the one employed in Grimshaw's theory of extended projection (Grimshaw 2005): for any branching non-terminal node the parser must encode whether it is part of the same extended projection as its left or its right daughter (Endocentricity).

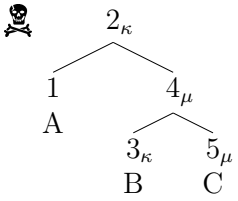
Notational conventions: (i) Greek subscripts indicate nodes belonging to the same extended projection; (ii) nodes are numbered; (iii) lexical material is represented as capitals below terminal nodes:

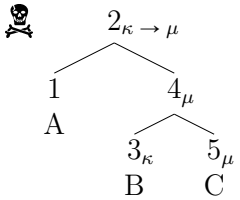
(30)


Consider two extensions of (30). As the parser fixes dominance but not immediate dominance relations, introduction of nodes between a mother and its daughter is information-preserving as long as headedness remains unaffected:

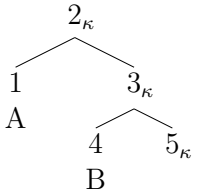
(31)


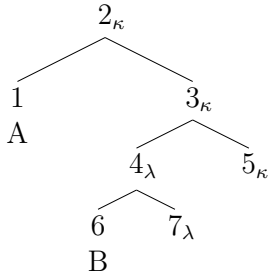
However, it is impossible to insert a node between the root and B that is part of a different extended projection. This violation of Endocentricity cannot be repaired by changing the index on 2 to μ, as that violates Monotonicity.

(32) a.


b.


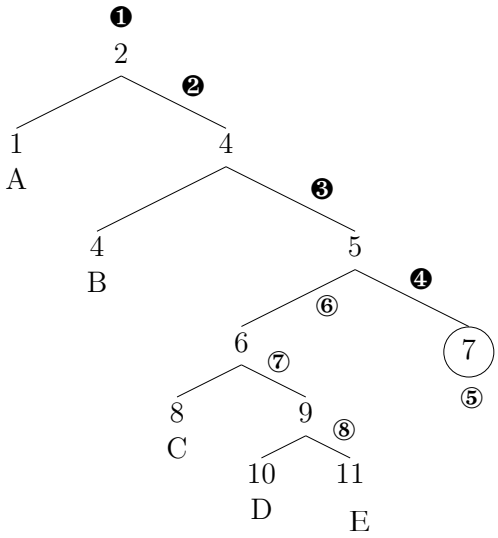
Suppose that neither A nor B can appear in the other's extended projection. In that case, an initial input A-B must be parsed as in (33a) or (33b), where 5 and 7 are empty nodes (instantiations of **e**).

(33) a.


b.


Empty nodes are predictions, and therefore subject to the LEC. The LEC facilites efficient parsing. In (34), for example, it reduces the possible attachment sites for an incoming category from eight to four:

(34)



This benefit does not come for free, as we have seen. We therefore assume that the parser only generating empty nodes when forced to do so.

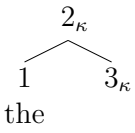
The procedure for parsing movement dependencies is the following:

(35) a. Identify the moved category.
b. Buffer it (that is, suspend work on it and store it).
c. Insert the copy.
d. Continue work on the constituent (if need be).

Several considerations favor a filler-driven strategy. (i) Traces have no phonological content. A gap-driven strategy must therefore rely on indirect evidence for gaps. However, such evidence is absent with adjunct movement or movement of optional arguments.

Moreover, there are many circumstances in which the parser postulates an empty node but should not start a search for a filler:

(36)



Phillips and Wagers (2007) give an overview of the substantial empirical support for the filler-driven strategy. To give one example, Stowe 1986 finds increased reading times for *us* in (37a) as compared to (37b) (the 'filled-gap effect'). This can be understood under the filler-driven strategy, but the gap-driven strategy has nothing to say about this.

(37) a. My brother wanted to know who Ruth will bring **us** home to at Christmas.
b. My brother wanted to know if Ruth will bring **us** home to Mom at Christmas.

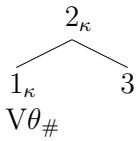
2. Parsing heavy-XP shift

We can now be more precise about why heavy-XP shift of the object is sensitive to obligatory transitivity.

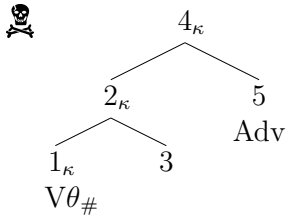
(38) a. John devoured t_1 yesterday [the food that his brother prepared]$_1$
b. John ate t_1 yesterday [the food that his brother prepared]$_1$

First consider *devoured yesterday the food that his brother prepared* (see (38a)). *Devour* requires an object:

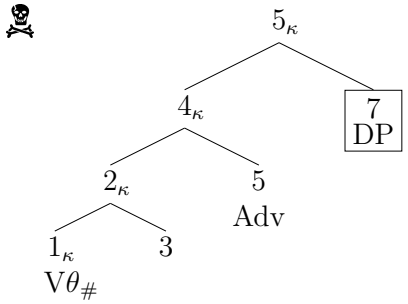
(39)



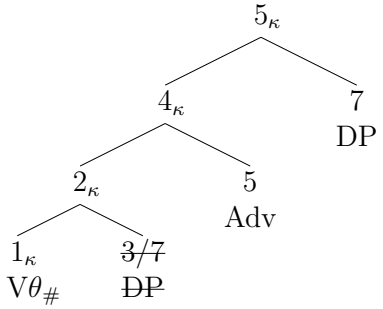
This adverbial *yesterday* cannot fill the complement position of the verb. Hence, the empty node is pushed away from the leading edge:

(40) 

The material following the adverbial is the onset of a DP. Once a position for this DP has been created, the parser can hypothesize that it has moved and start the search for a position in which to insert a trace:

(41) 

A scan of the parser's left context will easily identify an insertion site for the trace of the shifted object, namely the empty node adjacent to the verb:

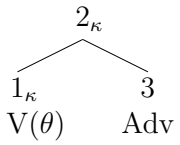
(42) 

Thus, heavy-XP shift of obligatory objects leads to difficulties in the parsing of material between the verb and the shifted DP and a speeding up once the DP has been encountered.

We now turn to the parsing of the substring *ate yesterday the food that his brother prepared* in (38b). We assume that optionally transitive verbs start out without an internal theta role, but may acquire one if necessary. Notation: '(θ).'

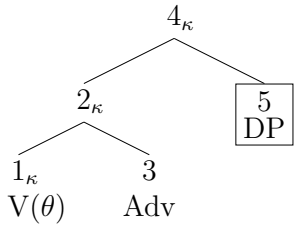
(43) 1_κ
V(θ)

The incoming adverbial can be accommodated without triggering a violation of the LEC:

(44) 

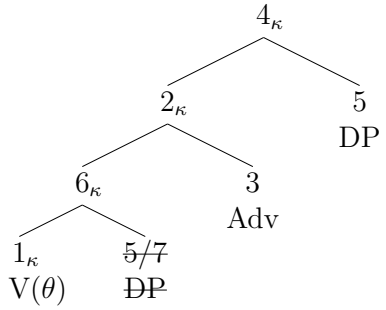
When the parser encounters the shifted DP, it will identify it as a moved category and start the search for a position to insert a trace.

(45)



Creation of this position involves postulation of nodes 6 and 7 in (46). This does not clash with any of the statements about dominance, precedence, or headedness that characterize (45). As soon as the empty position is created, it will be filled by a copy of DP, so that at no stage of the process the LEC is violated.

(46)

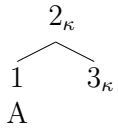


We thus have an explanation for the fact that heavy-XP shift of optional arguments does not cause parsing difficulties. In addition, we can understand why integration of the filler in (45) and (46) should be slower than in the obligatory condition. The absence of a pre-fabricated gap means that the search will not be aided by the existence of an unfilled position and that a position must be newly created.

3. Parsing movement of (constituents containing) the noun

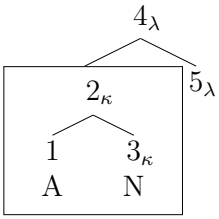
As an illustration, we consider two orders: A-N-Dem-Num and *Num-Dem-N-A. The first step in parsing the A-N-Dem-Num string involves providing a position for the adjective, which requires the postulation of an empty node:

(47)



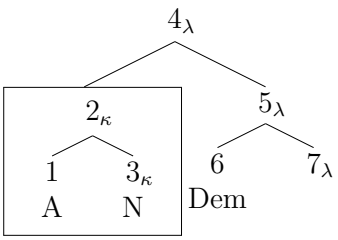
At the next step, N inserted in the empty node. On the assumption that the structure of the extended nominal projection always unfolds completely [A N] can be identified as a filler at this stage:

(48)

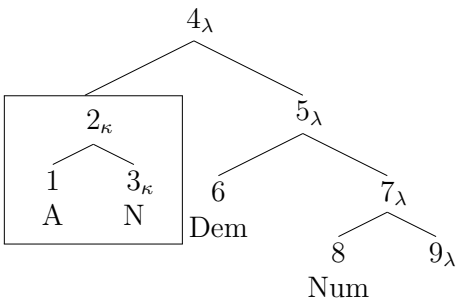


Subsequently, the demonstrative and the numeral are integrated, preserving the prediction that a trace is to follow:

(49) a.

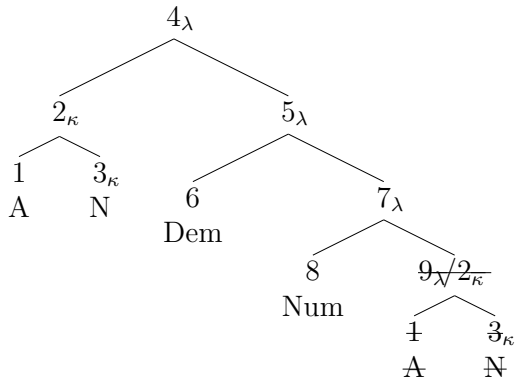


b.



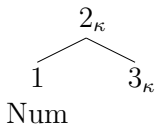
Finally, a trace (a copy of [A N]) is inserted for the empty node, leading to a full parse of the structure:

(50)



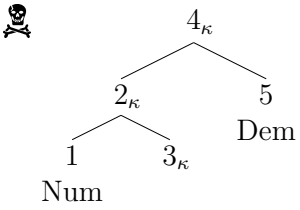
We next consider the mirror image of this order, *Num-Dem-N-A, which is unattested and would have to be derived by rightward movement of the [N A] constituent. The first category that the parser encounters is the numeral. On the assumption that the numeral is licensed in the extended projection of a noun, an empty node is predicted:

(51)



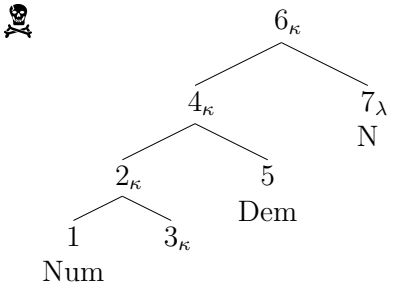
The demonstrative must be integrated in the nominal projection higher than the numeral. This means that the the empty node must be pushed away from the leading edge:

(52)

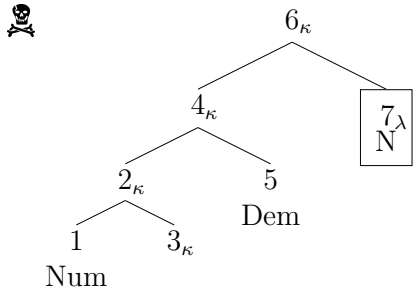


The next element is the noun, which can be integrated straightforwardly as having moved:

(53) a.

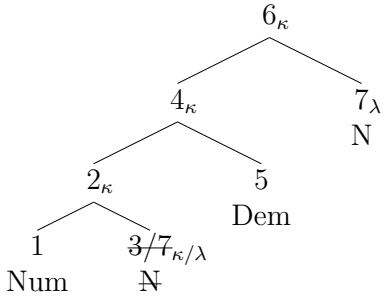


b.



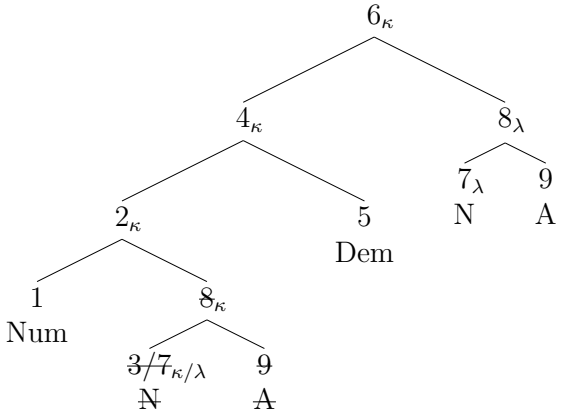
The movement dependency can be resolved by inserting a copy in empty node 3, so that compliance with the LEC is re-established:

(54)



Finally, the parser encounters the adjective, which would have to be integrated as in (55).

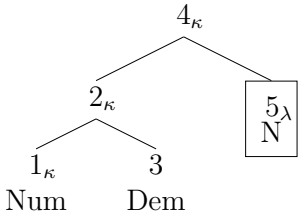
(55)



Thus, rightward movement of the noun-adjective combination in the Num-Dem-N-A order triggers a violation of the LEC.

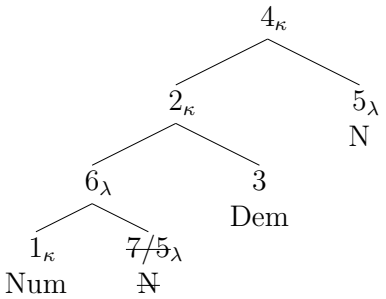
A crucial assumption in the derivation of this result is that the numeral must be licensed in the extended projection of the noun (which forces the parser to posit an empty node (node *3* in (51)). What happens if we allow the numeral to act as the head of an extended nominal projection? After encountering the the demonstrative and the noun, the parser ends up with the representation in (56).

(56)

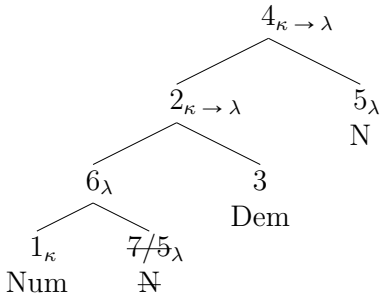


Resolution of the movement dependency leads to problems. As the noun heads the extended projection, it must percolate its index. Leaving the subscripts of nodes 2 and 4 intact leads to a violation of Endocentricity, but assigning nodes 2 and 4 a new subscript requires retraction of information and therefore violates Monotonicity:

(57) a. ☠



b. ☠



Thus, rightward movement of the noun within the noun phrase necessarily leads to parsing difficulties.

2. Projection and Dependencies

Inclusiveness, the main principle governing projection, if generalized beyond categorial features, can explain why syntactic dependencies require c-command. The account has implications for the analysis of movement that are not trivial but may help us dissolve the paradox resulting from Barss's generalization and the fact that remnant movement permits syntactic reconstruction.

This section is based on Neeleman, A., and H. van de Koot (2002). The Configurational Matrix. *Linguistic Inquiry* 33: 529-574, and Neeleman, A., and H. van de Koot. (2010). A Local Encoding of Syntactic Dependencies and its Consequences for the Theory of Movement. *Syntax* 13: 331-372.

2.1 Inclusiveness and projection

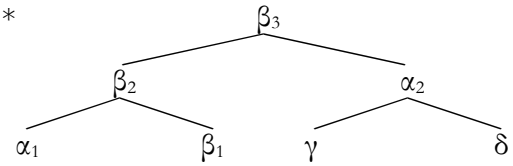
My starting point is a version of Inclusiveness that applies to each subtree (cf. Chomsky 1995: 228):

(1) *Inclusiveness*
The syntactic properties of a nonterminal node are fully recoverable from its daughters; the syntactic properties of a terminal node are fully recoverable through a pointer.

For now, a 'pointer' can be read as a pointer to a lexical entry from which the properties of a terminal node are copied.

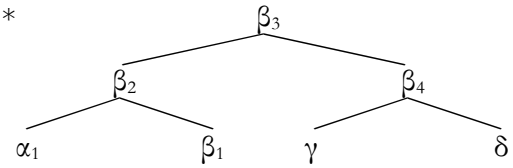
Inclusiveness restricts projection of categorial features. (i) It prevents sideward projection. Thus, β_1-β_2-β_3 in (2) is a well-formed projection, but α_1-α_2 is not (as α_1 is not a daughter of α_2).

(2) *



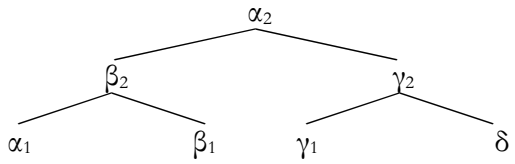
(ii) Inclusiveness rules out downward projection of categorial features. In (3), β_1-β_2-β_3-β_4 violates Inclusiveness because β_3 is not a daughter of β_4.

(3) *



(iii) Inclusiveness rules out discontinuous projection. In (4), α_1-α_2 is not a grammatical projection because α_1 is not a daughter of α_2.

(4) *



In combination with the standard assumption that each node must have a unique set of categorial features, this captures core properties of categorial projection.

2.2 Inclusiveness and dependencies

Certain dependencies are subject to a c-command requirement (Reinhart 1983). This is illustrated in (5) and (6) for secondary predication and anaphoric binding, respectively.

(5) a. Suzan heeft Ryo dronken aangesproken.
Suzan has Ryo drunk addressed
b. *Suzan's moeder heeft Ryo dronken aangesproken.
Suzan's mother has Ryo drunk addressed

c. Ik heb Ryo dronken naar huis gebracht.
I have Ryo drunk to home brought

d. *Ik heb met Ryo dronken gesproken.
I have with Ryo drunk spoken

(6) a. Ryo heeft zichzelf uitgenodigd.
Ryo has REFL-self invited

b. *Ryo's moeder heeft zichzelf uitgenodigd.
Ryo's mother has REFL-self invited

c. Ik heb Ryo met zichzelf geconfronteerd.
I have Ryo with REFL-self confronted

d. *Ik heb met Ryo over zichzelf gesproken.
I have with Ryo about REFL-self spoken

The c-command restriction has been argued to hold of movement, predication, thematic selection, anaphoric binding, and obligatory control. It is not a condition that holds of linguistic relations in general. For example, it does not hold of coreference (as illustrated in (7)). Neither does it hold of variable binding or non-obligatory control (Williams 1980, Reuland 2011, Barker 2012). The view I adopt here (following Koster 1987 and others) is that c-command is a hallmark of dependencies that are encoded syntactically.

(7) a. Ryo heeft z'n vader uitgenodigd.
Ryo has his father invited

b. Ryo's moeder heeft z'n vader uitgenodigd.
Ryo's mother has his father invited

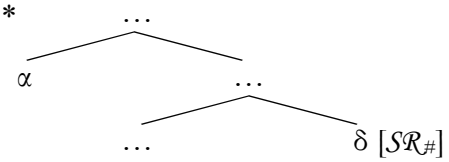
c. Ik heb Ryo met z'n vader geconfronteerd.
I have Ryo with his father confronted

d. Ik heb met Ryo over z'n vader gesproken.
I have with Ryo about his father spoken

Although the c-command requirement is a familiar one, it is not obvious why it should hold. I propose that an answer can be found in the unification of phrase structure theory and the theory of syntactic dependencies. I assume that a dependent category (an anaphor, a secondary predicate, etc.) is a dependent category because it contains a selectional requirement, that is, a feature that insists on the presence of a specific category (a binder, a subject, etc.) in its environment. When a selectional requirement finds what it is looking for, it is satisfied (marked by a subscript #).

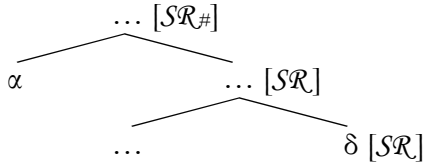
Consider the tree in (8), where δ is a syntactic dependent, α its antecedent and *SR* the selectional requirement that connects the two. In (8), *SR* is satisfied by α in a standard chain-like fashion.

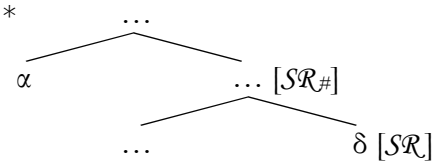
(8) *



The relation between α and δ determines properties of δ, because *SR* in δ cannot be satisfied again. Thus, the diacritic '#' represents a syntactic property of the node in which it occurs. Given that δ does not dominate α, the fact that *SR* in (8) carries '#' cannot be recovered from material internal to δ. Hence, a chain-like encoding of syntactic dependencies violates Inclusiveness.

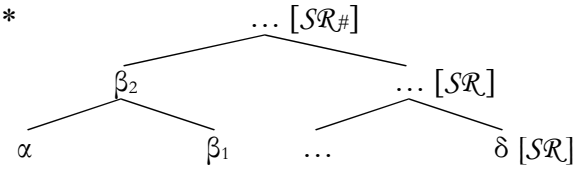
The logic of the problem dictates solution. Only if *SR* is copied upward recursively to the node that immediately dominates α, as in (9a), can it be satisfied without violation of Inclusiveness. If no node in the path from δ to α is skipped, upward copying is licit, because each new copy on a node can be recovered from its daughters. Downward satisfaction of *SR* also obeys Inclusiveness, because the element that determines the status of *SR* is a daughter of the node that contains *SR*. Notice that satisfaction under sisterhood, as in (9b), still violates Inclusiveness: the fact that *SR* is satisfied cannot be recovered from the daughters of the node that hosts it.

(9) a. 

b. * 

Thus, Inclusiveness forces a decomposition of syntactic dependencies into two primitive operations: the upward copying and downward satisfaction of a selectional requirement. Two key properties of syntactic dependencies follow from this.

- Dependencies may in principle span arbitrarily large distances. This is because a copied selectional requirement can itself be copied. As a result, the path along which a selectional requirement travels up the tree can in principle be indefinitely long (modulo locality).
- The relation between the antecedent and the highest node that contains $\mathcal{SR}$ is extremely local: the structure in (10), where β_2 intervenes between α and the node containing $\mathcal{SR}$, is ruled out, as it violates inclusiveness.

(10) * 

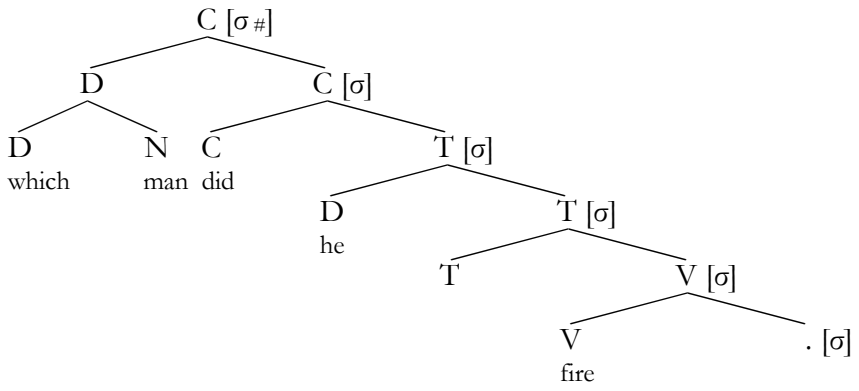
Thus, Inclusiveness does not only regulate projection, but it also explains why a dependent must be c-commanded by its antecedent.

One clarification is in order. The feature that satisfies a selectional requirement can itself have been copied. Given that this is possible, can it really be guaranteed that the antecedent occupies a c-commanding position? For example, could the feature of α that satisfies $\mathcal{SR}$ in (10) be copied to β and satisfy $\mathcal{SR}$ there? This problem is only apparent. Copied features become properties of the node they are copied to. Therefore, the relevant feature can only be copied to β if it is taken to characterize this node, in which case $\mathcal{SR}$ is in fact satisfied by β.

Of course, copied selectional requirements also become properties of the node they are copied to. However, whereas features indicate what a node *is*, selectional requirements indicate what a node *asks for*. In many cases, a node cannot be two things at the same time (say a verb and a preposition). But it can be one thing (say a verb) and ask for another (say a prepositional complement). Thus, copying of selectional requirements is less restricted than copying of features.

2.3 Movement

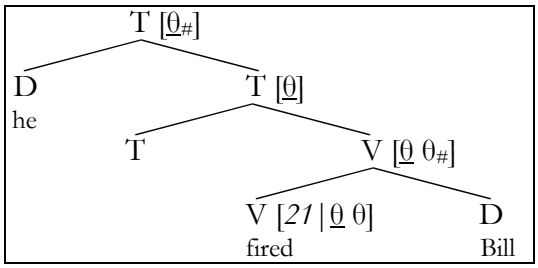
The above conception of syntactic dependencies implies that movement must be mediated by a selectional requirement introduced by a trace, copied up the tree, and satisfied by the trace's antecedent (here written as σ given that is reminiscent to the slash feature of HPSG):

(11) 

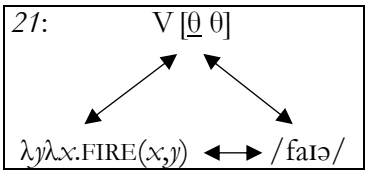
The selectional requirement σ introduced by the trace must be recoverable from the lexicon, suggesting that traces are lexical items. This has several consequences. To begin with, movement is known to reconstruct for syntactic and interpretive properties of the moved category and at first sight this seems to rule out an analysis of traces as simplex items stored in the lexicon.

How do other terminal nodes acquire their content? The notion of lexical insertion is a familiar metaphor, but not very helpful. Instead, the properties of a terminal node must match those of a lexical item. We assume that a syntactic terminal contains a pointer that identifies the lexical entry in question (Jackendoff 1996):

(12)



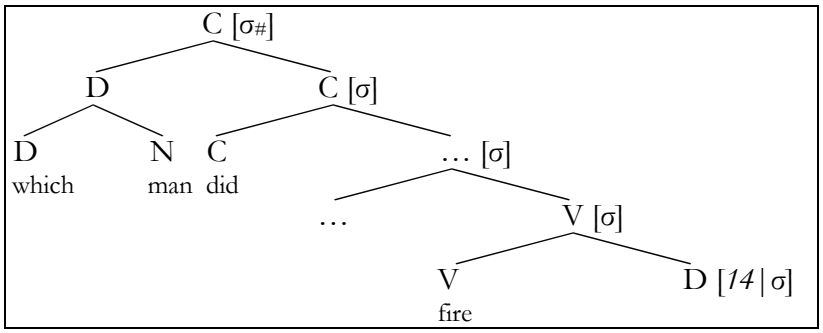
Syntax



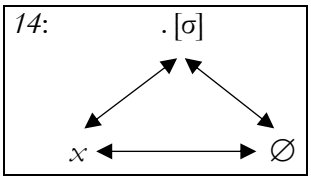
Lexicon

Given that matching can relate information contained in different representations, there can be no structural requirements on the nodes it relates. Therefore, nothing stands in the way of a terminal containing a pointer that relates it to another node in the same representation. Suppose that [μ] is a syntactic pointer. Then, the analysis of (11) can be improved as in (13)

(13)



Syntax



Lexicon

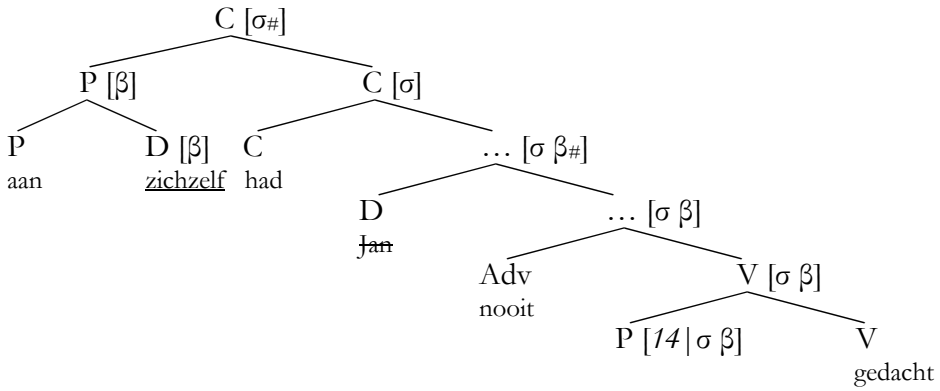
[σ] is licensed through a pointer to lexical entry 14 (the $\bar{A}$-trace). By being copied and satisfied in the usual way, it identifies a source for 'internal matching' of the remaining properties of the terminal. In the case at hand, it licenses a specification of the trace as a DP. Thus, traces are not copies of their antecedents, but copies of the top node of their antecedents.

This is enough to explain syntactic reconstruction effects, simply because on the proposed encoding of dependencies, the top node of the moved category contains more information than in alternative accounts. Consider the Dutch example in (14), where an anaphor is bound under reconstruction.

(14) [PP Aan zichzelf] had Jan nooit t_{PP} gedacht.
on self had John never thought
'Of himself John had never thought.'

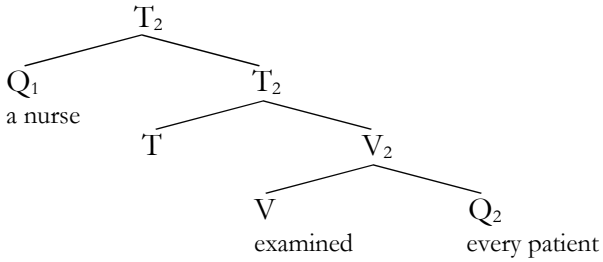
A partial structure corresponding to (14) is given below. The reflexive introduces the selectional requirement [β] responsible for anaphoric binding. The root of the moved category is formed by copying [β] from the reflexive and a set of categorial features from the preposition. The resulting node satisfies [μ], so that the trace is also a prepositional category carrying [β]. The instance of [β] on the trace undergoes copying in the usual way until it is satisfied by the subject.

(15)



Note that covert movement (quantifier raising) cannot be analyzed in the same way. Instead, I propose to use a mechanism proposed in Williams 1995. Each quantifier comes with a scope index. By default, a quantifier takes surface scope, but it can extend its scope by percolating its index (in line with Inclusiveness). The highest node with the index is construed as the quantifier's scope. Thus, Q_2 takes scope over Q_1 in (16).

(16)



2.4 Barss's paradox

The theory that traces are copies of the top node of their antecedents but have no internal structure can be used to explain a riddle in the theory of reconstruction. On the one hand, remnant movement allows reconstruction for syntactic properties, as illustrated in (17).

(17) [How likely t_1 to perjure <u>himself</u>]$_2$ does [<u>every politician</u>]$_1$ [seem [t_1 to be t_2]]?

On the other hand, it does not allow reconstruction for scope (an observation known as Barss's generalization). In order to see this, first consider (18), where the existential quantifier *some young lady* can be interpreted in the scope of the universal quantifier *every senator* (May 1977, Hornstein 1995, Fox 1999, and Lebeaux 2009).

(18) [Some young lady]$_1$ [seems (to Mary) [t_1 to be likely t_1 to dance with every senator]].
($\exists > \forall$; $\forall > \exists$)

This interpretation could result from reconstruction of the existential or from long-distance quantifier raising of the universal. There are several arguments that go against the latter option:

(i) If the universal could undergo long quantifier raising, it should be able to take scope over *likely*. This is not the case, however (May 1977). Consider an example like *Mary is likely to dance with every senator*. Given a sufficiently large number of senators, $\forall >$ *likely* would imply that it is *unlikely* that Mary will dance with every senator (if every senator is 90% likely to dance with Mary and there are 25 senators at the party, the chance that *every* senator will dance with Mary is 0.9^{25}, which equals 7%). That seems the wrong outcome.

(ii) Examples like (19) are unambiguous (Aoun 1982, Hornstein 1995, Fox 1999 and Lebeaux 2009). This remains unaccounted for if inverse scope results from long quantifier raising, but follows from the reconstruction account, because the fact that *some young lady* enters into a binding relation in its surface position blocks quantifier lowering.

(19) [Some young lady]$_1$ [seems to herself$_1$/her$_1$ companion [t_1 to be likely t_1 to dance with every senator]].
($\exists > \forall$; $*\forall > \exists$)

(iii) Lebeaux (2009) observes that in examples like (20) the existential cannot depend on the universal. Again, this is unexpected if inverse scope results from long quantifier raising but follows on the reconstruction account.

(20) Mary$_1$ [seems to some young lady [t_1 to be likely t_1 to dance with every senator]].
($\exists > \forall$; $*\forall > \exists$)

The key observation for our present purposes is that scope reconstruction of the existential becomes unavailable once the AP containing its trace undergoes *wh*-movement:

(21) [How likely t_1 to dance)with every senator]$_2$ does [some young lady]$_1$ [seem [t_1 to be t_2]]?
($\exists > \forall$; $*\forall > \exists$)

The observation that structures like (21) are unambiguous goes back to Barss 1986. Since Sauerland and Elbourne 2002, it is referred to as Barss's Generalization. My formulation is given in (22).

(22) *Barss's Generalization*: Scope reconstruction of a moved quantifier into a constituent X is blocked if the quantifier does not c-command X.

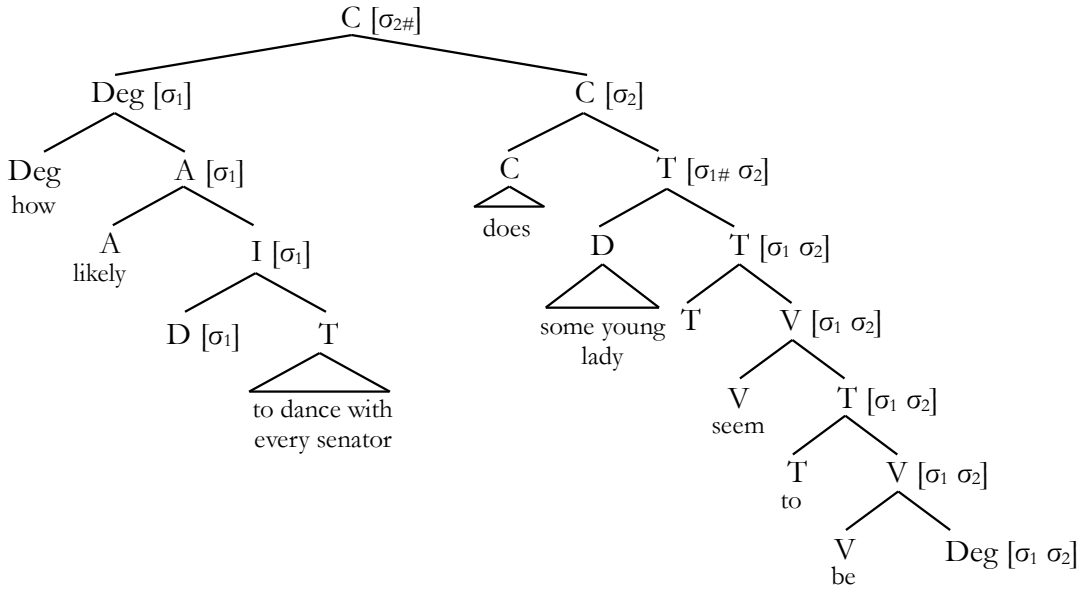
Semantic reconstruction must be distinguished from syntactic reconstruction (see Lechner 1998, 2013, 2019, Sternefeld 2000, Poole and Keine 2023). Fox (1999), looking at scope reconstruction and related phenomena, has suggested "that one can construct an argument for a [reconstruction site] in every maximal projection" (Fox 1999:175 n. 32). Within the current framework, this suggests the following rule:

(23) Let σ be a selectional requirement that encodes movement and let α be the category that satisfies it.
   a. The initial scopal domain of a is the node in which σ is satisfied.
   b. The scopal domain of α can be narrowed from n_1 to n_2 if n_1 and n_2 contain σ and n_1 immediately dominates n_2.

Now consider (24), where the remnant-creating movement is encoded by σ_1 and the remnant-movement by σ_2. The remnant-creating movement is licit because it relies on syntactic reconstruction. The relevant selectional requirement (σ_1) is present on the top node of the fronted category and therefore also present in the trace associated with that category.

The rule in (23) states that the scope of a moved category may be successively narrowed along the path of movement. The fact that the $\overline{\text{A}}$-movement in (24) removes part of the path of σ_1 from the structure dominated by $\sigma_{1\#}$ therefore implies that the number of scopal reconstruction sites is reduced as well. But this means that the existential cannot interact with the universal, thus explaining Barss's generalization. Ultimately, then, Barss's generalization falls out from the assumption that traces do not have internal structure.

(24)



Our account does not predict that either the raised subject or the remnant category must take surface scope: scope reconstruction is possible along the path of movement accessible to the rule in (23). Indeed, the raised subject can contain a variable bound under reconstruction (see (25a)), and so can the remnant (see (25b)).

(25) a. [How likely t_1 to pass the exam]$_2$ does [<u>his</u> weakest student]$_1$ appear to be t_2 to [<u>every professor</u>]?

b. [How likely t_1 to steal <u>his</u> favourite toy]$_2$ does Fido$_1$ appear to be t_2 to [<u>every boy</u>]?

2.5. Summary

- Inclusiveness subsumes two core properties of phrase structure theory: Endocentricity and Locality of Projection.
- It also forces a decomposition of syntactic dependencies into (potentially recursive) upward copying and local downward satisfaction of a selectional requirement. This, in turn, explains why syntactic dependencies require c-command.
- The proposal implies that traces must be terminals, and copies of the top node of their antecedent, rather than full copies. This in turn can explain why Barss's generalization should hold.
- Merge cannot explain the c-command requirement on movement chains.

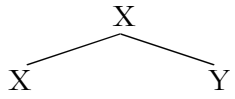
3. Subordination and Binary Branching

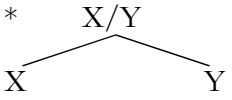
Phrase structure theory requires that subordination is licensed. The licensing system explains why subordination requires binary branching. It also predicts that structures that do not involve subordination can be flat (*n*-ary branching). This turns out to be true of coordination.

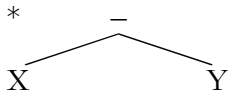
This section is based on Neeleman, A., J. Philip, M. Tanaka and H. van de Koot (2023). Subordination and Binary Branching. *Syntax* 26: 41–84.

3.1 Introduction

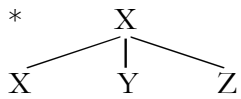
This paper is concerned with two constraints that impose asymmetry on syntactic representations. The first is *endocentricity*, which is in effect a ban on categories with multiple heads and categories without a head:

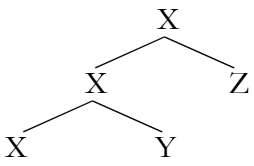
(1) a. 

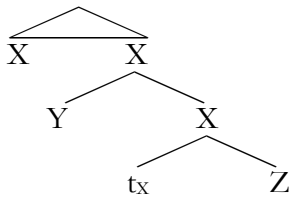
b. * 

c. * 

The second constraint requires that syntactic structures are *binary branching* (Kayne 1984), which imposes a structural asymmetry between dependents of a head:

(2) a. * 

b. 

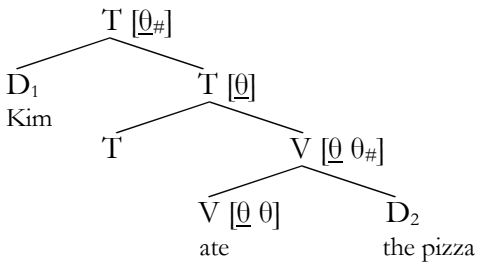
c. 

Endocentricity and the binary branching constraint are deeply engrained in generative grammar. In Minimalism, for example, merge takes two syntactic objects and delivers a new object that inherits its label from one of the input categories (Chomsky 1995). This set-up implies that binary branching has no exceptions but makes no predictions beyond this. We discuss a theory in which subordination must be licensed (where subordination is one category being contained in the projection of another), and licensing requires binary branching. The system imposes restrictions beyond binary branching per se (e.g. it rules out movement to a θ-position) and predicts that any non-endocentric syntactic structures can be *n*-ary branching. We argue coordinate structures are symmetric and need not be binary branching.

3.2 Toward an account of the Binary Branching Constraint

3.2.1 The Generalized Licensing Criterion

Structure must be licensed. We argue that a generalized version of the θ-criterion connects subordination to grammatical dependencies. The proposal relies on an encoding of grammatical dependencies through a selectional requirement copied upward along a connected path of nodes until it immediately dominates the antecedent (Neeleman and Van de Koot 2002, 2010):

(3) 

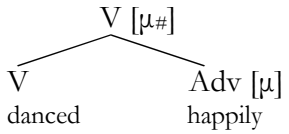
Part A of the proposed principle states that structure must be licensed when a category is subordinated to another category (Collins 1997, Chomsky 2000, Abels 2003, Wurmbrand 2014):

(4) *Generalized Licensing Criterion* (Part A; to be revised)
Subordination of Y to X requires a licensing relation between X and Y.

In (3), satisfaction of θ licenses subordination of D_2 to V, while subordination of D_1 to T is licensed by satisfaction of $\underline{\theta}$.

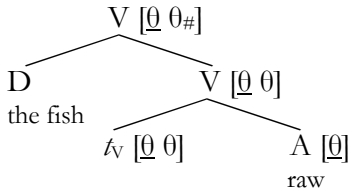
The selectional requirement that licenses subordination need not be introduced by the projecting node. Modifiers select the category they adjoin to (see Higginbotham 1985, Parsons 1990, Heim and Kratzer 1998, Maienborn 2001 and Ernst 2002). If this selectional relation is syntactically encoded through a selectional requirement μ, introduced by the modifier and satisfied by the projecting category, subordination of modifiers meets GLC-A.

(5)



Subordination may also be licensed through *identification* of selectional requirements (Higginbotham 1985). This is of particular relevance for secondary predication. As an example, consider *she ate the fish raw,* which contains a VP of the following shape.

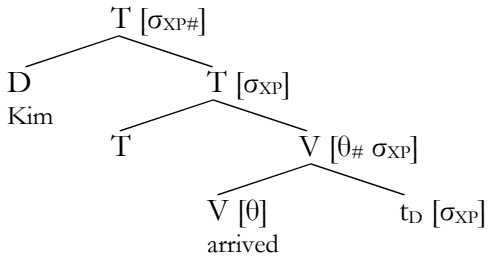
(6)



Here, subordination of DP to V' is licensed through satisfaction of a θ-role, and subordination of AP to V is licensed through θ-identification (which also reduces the total number of θ-roles). Hence, subordination of YP to X^n require the discharge of a selectional requirement. There are three modes of discharge: satisfaction by YP, satisfaction by X^n and identification.

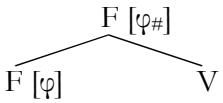
Not all syntactic dependencies can be used to license subordination. For example, it is not possible to project a position that hosts a DP if that DP merely binds an anaphor. Thus, the effects of the Generalized Licensing Criterion depend which selectional requirements have a licensing capacity. So far, we have assumed that θ and μ do. But subordination can also be licensed through movement:

(7)

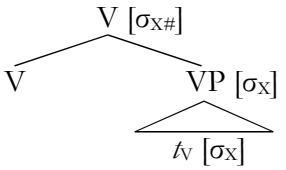


Two further licensing selectional requirements must be postulated to deal with functional structure. A base-generated functional head must select its complement (through φ.) Moved heads, we assume, are created through *self-attachment* (see Koeneman 2000, among many others). They connect to their trace through σ_x:

(8) a.



b.

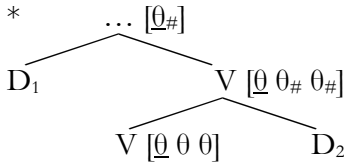


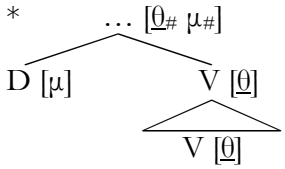
(9) *Generalized Licensing Criterion* (Part A; final version)
Subordination of Y to X requires a relation between X and Y that discharges a selectional requirement α (where $\alpha \in \{\theta, \mu, \sigma_{XP}, \varphi, \sigma_X\}$).

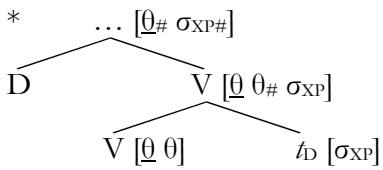
The standard θ-criterion banned multiple θ-role assignment to a single position. Within the current framework, this means that no node can be specified as [θ# θ#]. We propose this generalizes, too:

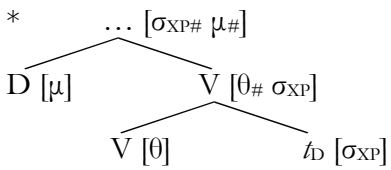
(10) *Generalized Licensing Criterion* (Part B; henceforth GLC-B)
No node created by subordination may be the locus of discharge of more than one selectional requirement taken from {θ, μ, σ_{XP}, φ, σ_X}.

Part B of the Generalized Licensing Criterion comprises fifteen cooccurrence restrictions that hold between satisfied selectional requirements. It would take us too far afield to discuss all fifteen. However, we will look in more detail at the six co-occurence restrictions that involve the phrasal dependencies encoded by θ, μ and σ_{XP}. The constraint in (8) rules out nodes specified as (i) [θ# θ#], (ii) [θ# μ#], (iii) [θ# σ_{XP}#], (iv) [μ# σ_{XP}#], (v) [μ# μ#], or (vi) [σ_{XP}# σ_{XP}#]. (The ban on movement to a θ-position is incompatible with the movement theory of control. See Hornstein & Polinsky 2010 Landau 2003, Landau & Bobaljik 2009 and Wood 2012 for discussion.)

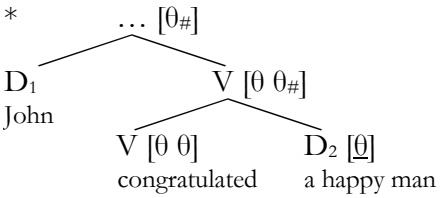
(11) a. *


b. *


c. *


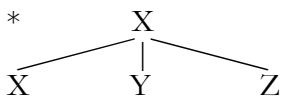
d. *


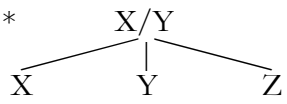
We further capture the fact that a DP argument cannot simultaneously act as depictive (even though there are DP depictives, as in *John left a happy man*):

(12) *


3.2.2 The source of the Binary Branching Constraint

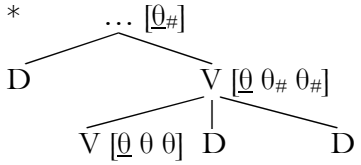
Crucially, the Generalized Licensing Criterion does not only account for the complementarity of certain syntactic functions, but also explains why subordination yields structures that are binary-branching (Kayne 1984).

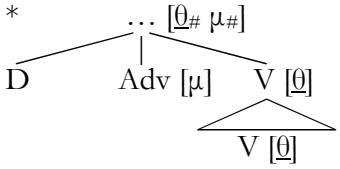
(13) a. *


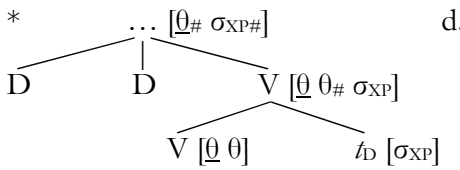
b. *


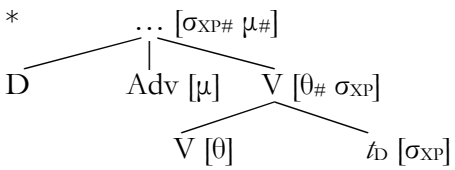
In both representations, there are two pairs consisting of a projecting and a non-projecting category. In (13a), these pairs are <X, Y> and <X, Z>; in (13b), they are <X, Z> and <Y, Z>. According to GCL-A, each pair must be licensed through the discharge of a selectional requirement. But according to GCL-B, no node created by subordination may be the locus of discharge of more than a single selectional requirement. Hence, (13a,b) are ruled out.

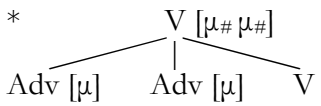
We give a range of example structures below.

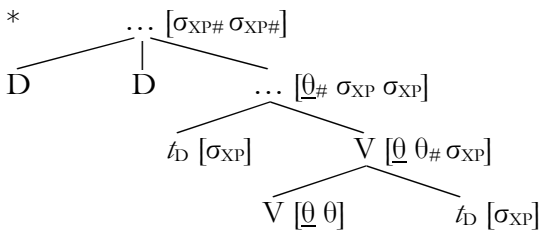
(14) a. *


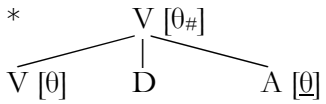
b. *


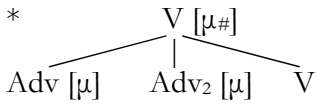
c. *


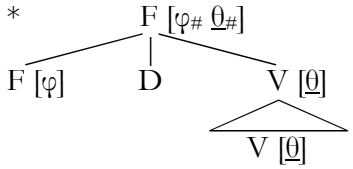
d. *


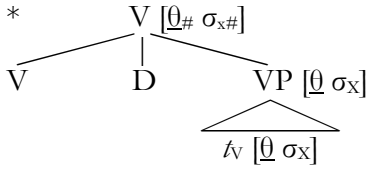
(15) a. * 

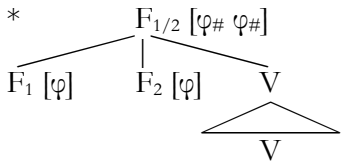
b. * 

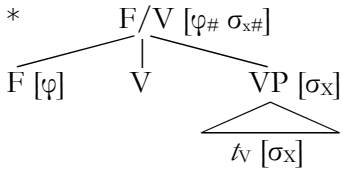
(16) a. * 

b. * 

(17) a. * 

b. * 

c. * 

d. * 

Thus, the same constraint that captures the complementarity of a range of syntactic functions also explains why subordination cannot create ternary-branching structures. This removes the necessity to state the ban on *n*-ary branching as a constraint on the input of merge. In addition, the proposal is open to falsification. We predict that there are no instances of subordination in which multiple or no selectional requirements are discharged.

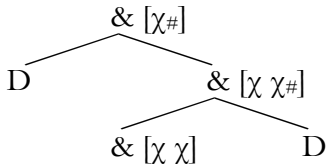
A counterexample of the latter type may come from expletives. There are three potential solutions. (i) An expletive could be a semantically bleached argument (see Chomsky 1981 on weather *it*, Bennis 1986 and Ruys 2010 on *it* expletives and Williams 1994, Hazout 2004, Hartmann 2008 and Van Craenenboeck 2020 on *there* expletives). (ii) An expletive could be a bleached modifier (see Bennis 1986 on Dutch *er* and Van Craenenbroeck 2020 on quasi-locational *there*). (iii) An expletive could be a bleached predicate (see Den Dikken 1995, Moro 1997 and Hartmann 2008).

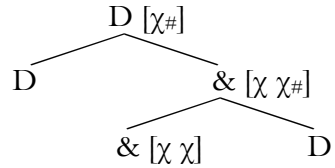
3.3 Coordination is not subordination

3.3.1 Coordination and the Generalized Licensing Criterion

If coordination is reduced to subordination, we end up with fully acceptable structures in which either GLC-A or GLC-B is violated. We first consider structures in which GLC-A is at stake.

Coordination cannot involve subordination of one conjunct to another, as conjuncts need not introduce any selectional requirements at all (as in *Thelma and Louise*). Hence, reducing coordination to subordination implies that the coordinator must introducing multiple selectional requirements, as in (18a) (Munn 1987, Kayne 1994, Zoerner 1995, Johannessen 1998 and De Vries 2005) or (18b) (Munn 1992, 1993, Bošković & Franks 2000, Hartmann 2000 and Zhang 2010).

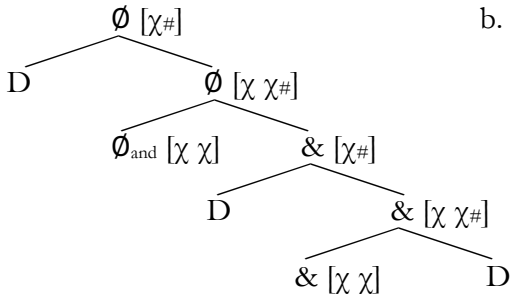
(18) a. 

b. 

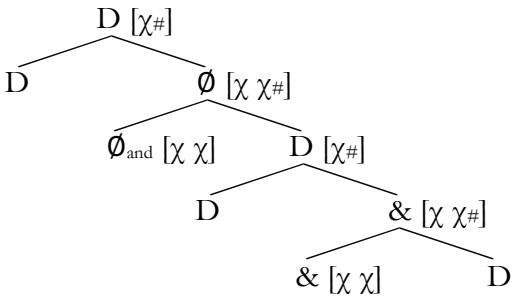
These analyses are self-defeating: given that the number of conjuncts is unbounded, the coordinator must contain an arbitrary number of selectional requirements, all but two of which are optional.

One could assume that a null coordinator is added for every conjunct beyond two (see Zoerner 1999 and De Vries 2005).

(19) a.



b.



How can we capture the distribution of overt and covert coordinators? First, if there is a single overt coordinator, it must attach to the final conjunct. Thus, [*Hal* [$Ø_{and}$ [*Thelma* [*and Louise*]]]] is grammatical, but *[*Hal* [*and* [*Thelma* [$Ø_{and}$ *Louise*]]]] is not. Second, there is no mixing of disjunctive and conjunctive coordinators (*Hal, Thelma and Louise* cannot mean *Hal or Thelma and Louise*, and *Hal, Thelma or Louise* cannot mean *Hal and Thelma or Louise*). Zoerner suggest that covert coordinators are in fact landing sites for LF raising of the overt coordinator. This captures the no-mixing restriction. However, it assumes that a coordinator's selectional requirements are 'reactivated' after each step of coordinator raising, which is equivalent to saying that a coordinator contains an arbitrary number of optional selectional requirements (in addition to two obligatory ones).

Other grammatical coordinate structures violate GLC-B if analyzed as subordination:

(20) Susan [$gave_V$ or $lent_V$] her two best friends all of her mother's books.

The only way to avoid this conclusion is to argue that examples like (20) involve phrasal coordination, with the surface form derived through some form of ellipsis:

(21) Susan [[$_{VP}$ gave ~~her two best friends all of her mother's books~~] or [$_{VP}$ lent her two best friends all of her mother's books]].

There are two arguments against reducing coordination of non-maximal categories to phrasal coordination plus ellipsis. (i) there may be a mismatch between the interpretations of the purported underlying and derived structures (see Borsley 2005:471 and Zhang 2010).

(22) a. Hobbs whistled and hummed a total of sixteen tunes.
b. Hobbs whistled ~~a total of sixteen tunes~~ and hummed a total of sixteen tunes.

(ii) The rules required to derive (20) from (21) are not well motivated. English does not have a general rule of backward ellipsis (see (23a)) and a right-node raising analysis is not feasible, given that examples like (20) do not have a right-node raising prosody and given that right-node raising of pronouns is awkward (Bresnan 1974:615), while examples like (23b) are unobjectionable.

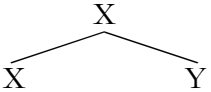
(23) a. *Susan [[$_{VP}$ gave ~~her two best friends~~ all of her father's records] or [$_{VP}$ lent her two best friends all of her mother's books]].
b. She [$gave_V$ or $lent_V$] it to him.

Thus, a subordination analysis of coordination leads to complications with both GLC-A and GLC-B.

3.3.2 Coordination as a symmetrical structure

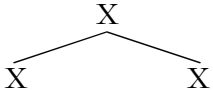
I analyze coordination as a symmetrical structure (Philip 2012). More specifically, the top node of a coordinate structure is a projection of two (or more) categories:

(24) a.



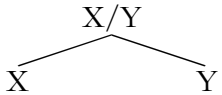
subordination

b.



coordination (like categories)

c.

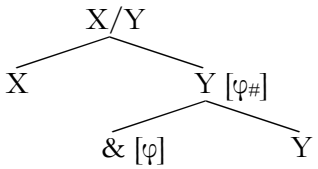


coordination (unlike categories)

(25) *Subordination*
A category Y is subordinated to a category X if the node immediately dominating Y carries X's categorial features, but not Y's categorial features.

The proposal does not immediately make clear what role the coordinator has. We turn to this in section 5. For now, we simply stipulate that a coordinator must be attached to the final conjunct and can optionally be attached to any medial conjuncts. Note that the coordinator must be a complete functor, passing up all properties of the category it selects.

(26)



The GLC applies to structures generated through subordination. In (22c), there is no subordination, and hence there is no demand for a selectional requirement to be discharged, nor for any discharge to be limited to a single selectional requirement. This permits coordination of arguments (without discharge) and coordination of transitive verbs (with multiple discharge).

The main empirical prediction of the proposal is known as Wasow's Generalization. The formulation in (27) is from (Pullum & Huddleston 2002: 1323).

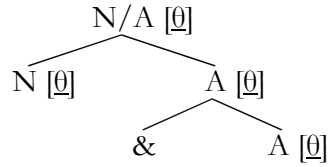
(27) *Wasow's Generalization*
If (and only if) in a given construction a constituent X can be replaced without change of function by a constituent Y, then it can also be replaced by a coordination of X and Y.

I first consider the effects of Wasow's Generalization for categorial features. It goes without saying that two categories of the same type can be coordinated in a position that admits that type. Given that *become* selects nominal and adjectival predicates, (28a) and (28b) are both grammatical. It is also possible to combine *become* with a mixed nominal/adjectival coordinate structure, as in (29) (on coordination of unlike categories, see Sag et al. 1985 and Bayer 1996).

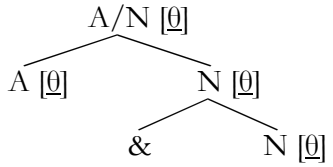
(28) a. Danny became [[$_{NP}$ a political radical] and [$_{NP}$ an angry man]].
b. Danny became [[$_{AP}$ very angry] and [$_{AP}$ very antisocial]].

(29) a. Danny became [[$_{NP}$ a political radical] and [$_{AP}$ very antisocial]].
b. Danny became [[$_{AP}$ very antisocial] and [$_{NP}$ a political radical]].

(30) a.



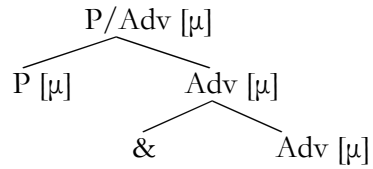
b.



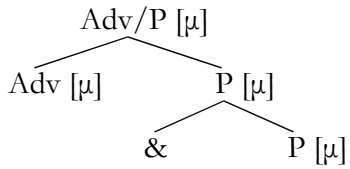
Similarly, modifiers of different category can be coordinated:

(31) a. We walked [[$_{PP}$ with great care] and [$_{AdvP}$ very slowly]].
b. We walked [[$_{AdvP}$ very slowly] and [$_{PP}$ with great care]].

(32) a.



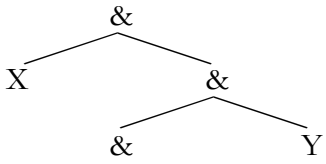
b.

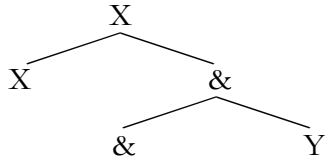


Become does not tolerate prepositional predicates, however, and hence it is not possible for *become* to combine with a coordination containing a prepositional conjunct:

(33) a. *Danny became [[$_{PP}$ under suspicion] and [$_{NP}$ a political radical]].
b. *Danny became [[$_{NP}$ a political radical] and [$_{PP}$ under suspicion]].
c. *Danny became [[$_{PP}$ under suspicion] and [$_{AP}$ very antisocial]].
d. *Danny became [[$_{AP}$ very antisocial] and [$_{PP}$ under suspicion]]

This effect is hard to understand if a coordinate structure is the projection of the coordinator or of the first conjunct. On neither analysis, it follows that the category of both conjuncts should matter:

(34) a. 

b. 

The effects of Wasow's Generalization extend to the selection of inflectional features:

(35) a. Hobbs turned out to like Rhodes and to hate Barnes.
b. *Hobbs turned out to like Rhodes and hating Barnes.
c. *Hobbs turned out liking Rhodes and to hate Barnes.

(36) a. Hobbs ended up liking Rhodes and hating Barnes.
b. *Hobbs ended up liking Rhodes and to hate Barnes.
c. *Hobbs ended up to like Rhodes and hating Barnes.

They also extend to the coordination of arguments (contra Bruening & Al Khalaf 2020). Patejuk & Przepiórkowski 2021 list numerous attested examples of unlike argument coordination, and the examples in (37) deteriorate sharply if the verbs are replaced by alternatives that tolerate only one of the conjoined categories as their internal argument, as in (38).

(37) a. John thought [CP that his budget would run out soon] and [PP about the difficulties this would cause].
b. Fiona noticed [CP that the numbers for the tax year did not add up] and [DP a range of other errors in Bill's accounting].
c. The strikers fought [DP the college's injustice] and [PP against those that kept it in place].

(38) a. *John noticed [CP that his budget would run out soon] and [PP about the difficulties this would cause].
b. *Fiona said [CP that the numbers for the tax year did not add up] and [DP a range of other errors in Bill's accounting].
c. *The strikers resisted [DP the college's injustice] and [PP against those that kept it in place].

Coordination of arguments does give rise to one tricky issue, though: there are grammatical examples in which a verb selects a conjunction of a DP and a CP, even though it does not take CP complements in isolation:

(39) You can depend on my assistant and that he will be on time.

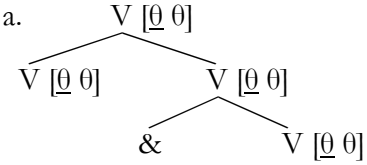
Bruening & Al Khalaf (2020) argue that examples of this type involve coordination of like categories, with the CP conjunct a DP headed by a silent noun. This proposal predicts (correctly, it seems) that violations of Wasow's Generalization are restricted to cases in which a CP is coordinated with a DP.

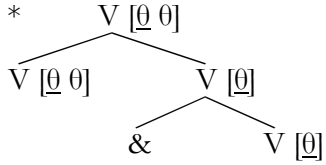
Further empirical consequences of the proposal depend on what is deemed to break the symmetry of a coordinate structure. We assume (40) holds. Recall that licensing selectional requirements license subordination. Hence (40) implies that conjuncts demand the same structural environment.

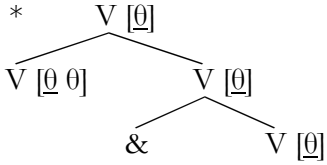
(40) In a coordinate structure, each conjunct must have the same set of unsatisfied licensing selectional requirements as its mother.

It follows that when two predicates (say, two verbs) are coordinated, they must have the same arity:

(41) a. I saw him [[buy] and [read]] a book.
b. *I saw him [[buy] and [sleep]] a book.

(42) a. 

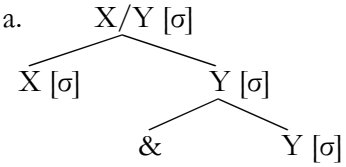
b. * 

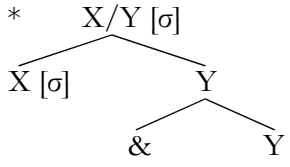
c. * 

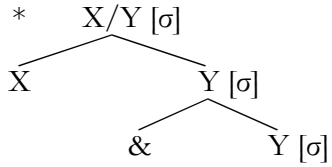
It also follows that movement out of coordinate structures is subject to the element constraint (the conjunct constraint remains unexplained).

(43) *Coordinate Structure Constraint* (Ross 1967, Grosu 1973, De Vries 2017)
  a. The Conjunct Constraint: Conjuncts cannot be moved.
  b. The Element Constraint: Movement out of a coordinate structure is possible only if the moved category binds a trace in each conjunct.

(44) a. The madrigals which$_1$ Henry [[sings t_1] and [listens to t_1]] are mostly Venetian.
  b. *The lute which$_1$ Henry [[plays t_1] and [sings madrigals]] is warped.
  c. *The madrigals which$_1$ Henry [[plays the lute] and [sings t_1]] sound lousy.

(45) a.


b. *


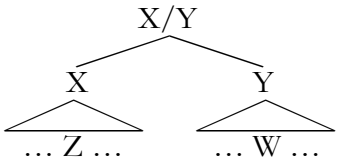
c. *


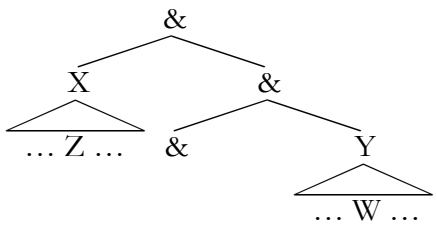
I will come back to apparent exceptions to the coordinate structure constraint.

In sum, at least some core properties of coordinate structures follow straightforwardly from a symmetrical analysis.

3.3.3 Coordination and scope

If coordinate structures are headed by a coordinator, the left conjunct c-commands and hence takes scope over the right conjunct. A symmetrical analysis implies that in a two-termed coordinate structure neither conjunct takes scope over the other, as neither is in a structurally higher position. Hence, in (46a) W cannot be scopally dependent on X and Z cannot be scopally dependent on Y. This is different on other accounts of coordination. In (46b), for example, W could be scopally dependent on X (though not Z on Y).

(46) a.


b.


How, then, to account for examples like *every man and his dog*? (Munn 1993). We adopt Progovac's (2000, 2003, a.o.) proposal that this requires quantifier raising of *every man* (here analyzed as percolation of a scope index). In general, a quantifier in a two-termed coordination may raise out of a left conjunct as long as it binds a variable in the right conjunct (see Rodman 1976, Ruys 1992 and Fox 2000):

(47) a. A soldier [[found every traitor] and [left unseen]]. (*∀>∃)
  b. A soldier [[found [every traitor]$_1$] and [shot him$_1$]]. (∀>∃)

(48) a. *Which student [[likes which professor] and [hates the dean]]? (on a pair-list reading)
  b. Which student [[likes which professor]$_1$ and [wants him$_1$ to be on his committee]]?

Notice that the element in the right conjunct must be a variable:

(49) a. A (different) student [[likes [every professor]] but [hates some TAs]].
  (*'[Every professor]$_1$ is such that a different student likes him$_1$ but hates some TAs.')
  b. A (different) student [[likes [every professor]$_1$] but [hates some of his$_1$ TAs.]]
  ('For [every professor]$_1$ there is a student that likes him$_1$ but hates some of his$_1$ TAs.')
  c. A (different) student [[likes [every professor]$_1$] but [wants him$_1$ to fire some TAs]].
  ('For [every prof.]$_1$ there is a student that likes him$_1$ but wants him$_1$ to fire some TAs.')

The data are as predicted by the symmetric analysis:

(50) a. Every man and a woman walked in. (*∀>∃)
  b. [Every man]$_1$ and his$_1$ wife walked in.

c. [Every man]$_1$ and a woman he$_1$ used to date walked in. (∀>∃)

(51) a. *Which priest united which refugee and three compatriots? (on a pair-list reading)
b. Which priest united [which refugee]$_1$ and his$_1$ family? (on a pair-list reading)

Progovac's proposal for *every man and his dog* receives further support from the example in (52):

(52) I remember [every corrupt politician]$_1$ and the false promises [the bastard]$_1$ made.

An epithet bound by a quantifier is subject to Principle C (Hornstein & Weinberg 1990:134). Hence, the example in (53), where the quantifier phrase uncontroversially c-commands the epithet, is ungrammatical:

(53) *[Every corrupt politician]$_1$ certainly likes the false promises that [the bastard]$_1$ makes.

Since (52) is grammatical, it follows that the left conjunct (*every corrupt politician*) does not c-command the right conjunct, but rather binds the epithet after quantifier raising.

3.3.4 Other asymmetries

The distribution of cases in a coordinate structure is not always symmetrical and there are situations in which the verb agrees with one conjunct but not the other ('unbalanced coordination', Johannessen 1998).

Unbalanced case. A recent paper by Weisser (2020) shows that in certain coordinate structures that appear asymmetric on the surface underlying case is in fact symmetric. For various other classes of data, Przepiórkowski (2021) argues that genuine mismatches in case can be observed, but these do not motivate a structural asymmetry between conjuncts. For example, Polish allows coordination of accusative and genitive DPs (if the latter have a partitive reading), but does not impose constraints on their respective order.

Unbalanced agreement. Agreement is coordinate structures comes in three patterns. (i) Resolution: the features of all conjuncts are input to a computation that derives a set of output features relevant for agreement. Resolution treats all conjuncts on a par and hence provides no evidence for syntactic asymmetry. (ii) Closest conjunct agreement: a predicate agrees with the conjunct linearly closest to it (Corbett 2006, Marušič et al. 2015, Nevins & Weisser 2019). Again, this pattern does not provide evidence for syntactic asymmetry. (iii) Distant conjunct agreement, as found in Slovenian (Marušič et al. 2015, a.o.): a predicate can agree with the first conjunct of a coordinate subject that precedes it, but not with the last conjunct of a coordinate subject that follows it. This pattern is typologically rare.

(54) a. Knjig-e in peres-a so se podražil-a/?podražil-e.
book-F.PL$_1$ and pen-N.PL$_2$ AUX.PL REFL got.dear-N.PL$_2$/got.dear-F.PL$_1$
'Books and pens have become more expensive.'
b. Podražil-a/*podražil-e so se peres-a in knjig-e.
got.dear-N.PL$_1$/got.dear-F.PL$_2$ AUX.PL REFL pen-N.PL$_1$ and book-F.PL$_2$

Our impression, based on work with three linguistically trained native speakers, is that the interpretation of the coordinate subject is an important factor. A coordination *A and B* can describe a set {A, B}, but it also allows a comitative reading *A and also B*. The comitative reading facilitates distant conjunct agreement if (in descriptive terms) A is foregrounded and B backgrounded:

(55) **Knjig-e** pa tudi peres-a so se podražil-e.
book-F.PL$_1$ and also pen-N.PL$_2$ AUX.PL REFL got.dear-F.PL$_1$

In context, comitative coordinate structures permit foregrounding of the second conjunct. Under those circumstances our informants found postverbal distant conjunct agreement unobjectionable:

(56) [Peresa so se resnično podražila.]
'Pens really got more expensive.'
Podražil-e so se ne samo peresa ampak tudi **knjig-e**.
got.dear-F.PL$_2$ AUX.PL REFL not only pen-N.PL$_1$ but also book-F.PL$_2$
'Not only pens but also books got more expensive.'

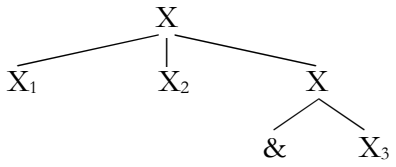
If so, it is not clear that distant conjunct agreement supports an asymmetric analysis of coordination after all.

3.4. *N*-ary branching coordinate structures

3.4.1 Flat and articulated coordinate structures

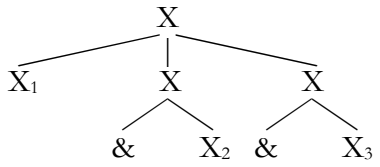
We now come to the crux of the paper. We have argued that subordination must meet the Generalized Licensing Criterion, and that this implies that subordination must be binary branching. We have also argued that coordinate structures do not involve subordination. Hence, coordinate structures can, but do not have to be binary branching:

(57) a.



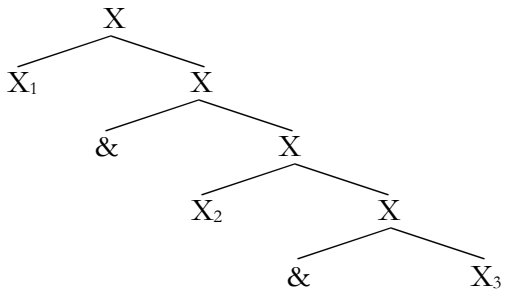
X_1 | X_2 | and X_3

b.



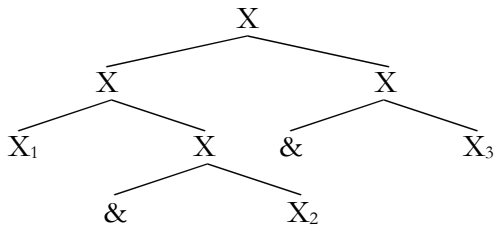
X_1 | and X_2 | and X_3

(58) a.



XP_1 ‖ and XP_2 | and XP_3

b.



XP_1 | and XP_2 ‖ and XP_3

In English coordinate structures, a coordinator must (normally) be attached to the final conjunct. Given that (58a) and (58b) have an articulated structure, a coordinator must hence appear between each pair of conjuncts, as indicated. The structure in (57) requires a single coordinator, attached to XP_3. Attachment of a coordinator to the medial conjunct is optional.

The embedding of one coordinate structure within another is reflected in the prosody (Zoerner 1995, Wagner 2010 and Kentner & Féry 2013):

(59) a. [Hal and Thelma and Louise]
b. [Hal 'n' Thelma 'n' Louise]
c. ??[Hal and Thelma 'n' Louise]
d. ??[Hal 'n' Thelma and Louise]

(60) a. [[Hal and Thelma] and Louise]
b. ??[[Hal 'n' Thelma] 'n' Louise]
c. *[[Hal and Thelma] 'n' Louise]
d. [[Hal 'n' Thelma] and Louise]

(61) a. [Hal and [Thelma and Louise]]
b. ??[Hal 'n' [Thelma 'n' Louise]]
c. [Hal and [Thelma 'n' Louise]]
d. *[Hal 'n' [Thelma and Louise]]

A reviewer suggests that the claim that monosyndetic multi-termed coordinate structures are *n*-ary branching runs foul of examples like (62a) and especially (62b):

(62) a. I lived in Berlin last year, and in Munich and in Paris.
b. I lived in Berlin last year, in Munich, and in Paris.

However, apparent extraposition out of coordinate structures results from partial ellipsis of a clausal conjunct (Johnson 2004, Chaves 2007 and Zhang 2010, a.o.):

(63) a. *I introduced Carrie to each other and Will.
b. [I introduced Carrie to each other] and [Will$_1$ ~~I introduced~~ t_1 ~~to each other~~].
c. I introduced [Carrie [$t_{\&DP}$]] to each other [and Will].

If the above in on the right track, (62a,b) should be analyzed as in (64a,b), respectively.

(64) a. [$_S$ I lived in Berlin last year], and [$_S$ [in Munich and in Paris]$_1$ ~~I lived~~ t_1 ~~last year~~]
b. [$_S$ I lived in Berlin last year], [$_S$ [in Munich] $_1$ ~~I lived~~ t_1 ~~last year~~], and [$_S$ [in Paris]$_2$ ~~I lived~~ t_2 ~~last year~~]

Further support for an ellipsis account comes from the distribution of coordinators in variants of (62a,b) in which only the final conjunct has shifted rightwards. If we were dealing with extraposition, there would be a single underlying coordinate structure, and hence a single coordinator preceding the final locational modifier should suffice, contrary to fact:

(65) a. *I lived in Berlin, in Munich last year, and in Paris.
b. I lived in Berlin and in Munich last year, and in Paris.

3.4.3 Literature review

If *Hal, Thelma and Louise* has a flat structure, it must denote a single three-termed coordination. It follows that it cannot mean the same as *Hal or Thelma and Louise*. Similarly, *Hal, Thelma or Louise* cannot be interpreted as *Hal and Thelma or Louise*.

Even when the no-mixing restriction is adhered to, the coordinate structures in (42) and (44) have different truth conditions (see Borsley 1994, 2005; Wagner 2010 and Winter 2006).

(66) [Tom and Dick] lifted the piano. (Borsley 1994:238)
a. 'Tom lifted the piano and Dick lifted the piano.' (distributive)
b. 'Tom and Dick together lifted the piano.' (collective)

(67) [Tom and Dick and Harry] lifted the piano. (Borsley 1994:238)
a. 'Tom lifted the piano, Dick lifted the piano, and Harry lifted the piano.' (distributive)
b. 'Tom, Dick and Harry together lifted the piano.' (collective)
c. 'Tom lifted the piano, and Dick and Harry together lifted the piano.' (mixed dist.-col.)
d. 'Tom and Dick together lifted the piano, and Harry lifted the piano.' (mixed col.-dist.)

(68) [Tom, Dick and Harry] lifted the piano. (Borsley 1994:239)
a. 'Tom lifted the piano, Dick lifted the piano, and Harry lifted the piano.' (distributive)
b. 'Tom, Dick and Harry together lifted the piano.' (collective)
c. *'Tom lifted the piano, and Dick and Harry together lifted the piano.' (mixed dist.-col.)
d. *'Tom and Dick together lifted the piano, and Harry lifted the piano.' (mixed col.-dist.)

In (68), the mixed readings are absent, because in a ternary-branching analysis no two conjuncts form a coordinate structure to the exclusion of the third.

Winter (2006) observes a similar effect with adverbials of alternation:

(69) a. John alternately feels [guilt and anger and hate] for his family. (Winter 2006:9)
b. % John alternately feels [guilt, anger and hate] for his family.

(70) a. two-state alternation: guilt/[anger and hate] - ✓(69a); *(69b)
b. two-state alternation: [guilt and anger]/hate - ✓(69a); *(69b)
c. three-state alternation: guilt/anger/hate - %(69a); %(69b)

The two state alternations are predicted to be absent in flat structures.

Borsley (2005) reports a related pattern for examples containing *respectively*. This adverb 'establishes a pairing between elements of two sets having the same cardinality' (Dalrymple & Kehler 1995:536 and references cited there):

(71) The two girls were seen by [Hobbs and Rhodes], respectively. (Borsley 2005:469)

The constrast between (72) and (73) is now predicted:

(72) The two girls were seen by [Hobbs and Rhodes and Barnes], respectively. (Borsley 2005:470)
a. 'Hobbs saw one girl, and Rhodes and Barnes saw the other.'
b. 'Hobbs and Rhodes saw one girl, and Barnes saw the other.'

(73) *The two girls were seen by [Hobbs, Rhodes and Barnes], respectively. (Borsley 2005:470)

A final data set confirming the same conclusion comes from the distribution of *both*, which can introduce a coordination whose cardinality is exactly two:

(74) a. [both [Tom [and Dick and Harry]]]
b. [[both [Tom and Dick]] and Harry]
c. [both [[Tom and Dick] and Harry]] (Borsley 1994:237)

(75) a. *both Tom, Dick and Harry (Borsley 1994:237)
b. *Tom, both Dick and Harry

In sum, in three-element conjunctions with a single coordinator neither the final, nor the initial two conjuncts entertain a distinct conjunctive relationship. This follows if such examples have a flat structure, as in (57a).

3.4.4 An argument from modification

If we were dealing with a binary-branching structure, either the first two conjuncts (in a left-branching structure), or the last two conjuncts (in a right-branching structure), form a constituent to the exclusion of the remaining conjunct. On the other hand, on the *n*-ary branching analysis in (42) no two conjuncts form a constituent to the exclusion of the other. These predictions can be tested by looking at the scope of modifiers.

In a binary coordinate structure, the presence of an initial results in ambiguity. In unambiguously flat multitermed coordinations, modification applies to the immediately following conjunct or to all conjuncts if the modifier is initial, but never to a non-singleton subset of conjuncts:

(76) Mary will buy [yellow pansies and tulips].
a. [yellow pansies] [and tulips]
b. [yellow [pansies and tulips]]

(77) Mary will buy [crocuses and yellow pansies and tulips].
a. crocuses and [yellow pansies] and tulips
b. crocuses and [yellow [pansies and tulips]]

(78) Mary will buy [yellow crocuses and pansies and tulips].
a. [yellow crocuses] and pansies and tulips
b. [yellow [crocuses and pansies]] and tulips
c. [yellow [crocuses and pansies and tulips]]

(79) Mary will buy [crocuses, yellow pansies and tulips].
a. crocuses, [yellow pansies] and tulips
b. *crocuses, [yellow [pansies and tulips]]

(80) Mary will buy [yellow crocuses, pansies and tulips].
a. [yellow crocuses] [pansies] [and tulips]
b. *[yellow [crocuses, pansies]] [and tulips]
c. [yellow [crocuses, pansies and tulips]]

The same pattern is found with adjuncts that follow the constituent they modify:

(81) The park was full of [dog-walkers and tourists and children on scooters].
a. dog-walkers and tourists and [children on scooters]
b. dog-walkers and [[tourists and children] on scooters]
c. [[dog-walkers and tourists and children] on scooters]

(82) The park was full of [dog-walkers, tourists and children on scooters].
a. dog-walkers, tourists and [children on scooters]
b. *dog-walkers, [[tourists and children] on scooters]
c. [[dog-walkers, tourists and children] on scooters]

(83) The park was full of [dog-walkers and tourists on scooters and children].
a. dog-walkers and [tourists on scooters] and children
b. [[dog-walkers and tourists] on scooters] and children

(84) The park was full of [dog-walkers, tourists on scooters and children].
  a. dog-walkers, [tourists on scooters] and children
  b. *[[dog-walkers, tourists] on scooters] and children

The examples in (81) and (83), where a coordinator occurs between each pair of conjuncts and hence the binary branching structures in (58) are possible, have a reading in which *on scooters* takes scope over the embedded two-termed coordination to its left, yielding the (b) reading. This reading is not available in (82) and (84), which contain only a single coordinator.

The pattern repeats itself with coordinated VPs modified by a manner adverbial: the (b) readings available in (85), (87) (89) and (91) drop out in (86), (87), (91) and (92).

(85) John hurriedly [got up] and [put on a suit] and [walked to the office].
  a. [hurriedly [got up]] and [put on a suit] and [walked to the office]
  b. [hurriedly [got up] and [put on a suit]] and [walked to the office]
  c. [hurriedly [got up] and [put on a suit] and [walked to the office]]

(86) John hurriedly [got up], [put on a suit] and [walked to the office].
  a. [hurriedly [got up]], [put on a suit] and [walked to the office]
  b. *[hurriedly [got up], [put on a suit]] and [walked to the office]
  c. [hurriedly [got up], [put on a suit] and [walked to the office]

(87) John [got up] and hurriedly [put on a suit] and [walked to the office].
  a. [got up] and [hurriedly [put on a suit]] and [walked to the office].
  b. [got up] and [hurriedly [put on a suit] and [walked to the office]].

(88) John [got up], hurriedly [put on a suit] and [walked to the office].
  a. [got up], [hurriedly [put on a suit]] and [walked to the office].
  b. *[got up], [hurriedly [put on a suit] and [walked to the office]].

(89) John [got up] and [put on a suit] and [walked to the office] in a hurry.
  a. [got up] and [put on a suit] and [[walked to the office] in a hurry]
  b. [got up] and [[put on a suit] and [walked to the office] in a hurry]
  c. [[got up] and [put on a suit] and [walked to the office] in a hurry]

(90) John [got up], [put on a suit] and [walked to the office] in a hurry.
  a. [got up], [put on a suit] and [[walked to the office] in a hurry]
  b. *[got up], [[put on a suit] and [walked to the office] in a hurry]
  c. [[got up], [put on a suit] and [walked to the office] in a hurry]

(91) John [got up] and [put on a suit] in a hurry and [walked to the office].
  a. [got up] and [[put on a suit] in a hurry] and [walked to the office]
  b. [[got up] and [put on a suit] in a hurry] and [walked to the office]

(92) John [got up], [put on a suit] in a hurry and [walked to the office].
  a. [got up], [[put on a suit] in a hurry] and [walked to the office]
  b. *[[got up], [put on a suit] in a hurry] and [walked to the office]

3.5 The status and distribution of coordinators

Symmetric coordination does not inherently require insertion of a coordinator. After all, the coordinator is not the head of the structure. Indeed, asyndetic coordination is common cross-linguistically, and perhaps available universally (Payne 1985 and Haspelmath 2007). English examples disjunction are given in (93). As noted (e.g. Büring and Hartmann 2015), asyndetic coordination gives an impression of incompleteness or open-endedness

(93) a. He had brought [gifts, flowers, chocolate, champagne], and yet he felt unwelcome.
  b. I found no more than [two, three] mistakes in your article.

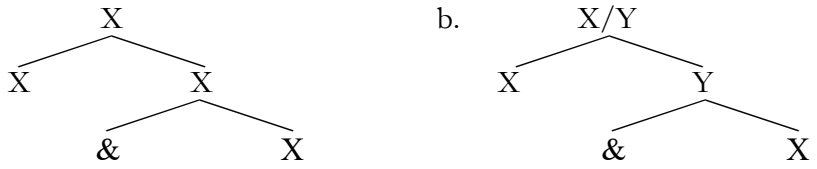
Thus, in English multi-termed coordinations, there may be zero, one or multiple coordinators. If there is one coordinator, it is attached to the final conjunct; if there are multiple coordinators these are attached to all non-initial conjuncts.

Coordinators are not attached in a completely flat structure (contra Dik 1968, Goodall 1987 and Muadz 1991). In (94), the coordinator cannot subordinate, or be subordinated to, the conjuncts. (This would lead to a structure subject to the GLC and that hence cannot be ternary branching.) But then the coordinator must itself be a conjunct – clearly an incoherent result.

(94) a. * X b. * X/Y

Therefore, the coordinator must be attached to a conjunct (Ross 1967, Zwart 2009 and Philip 2012):

(95) a. b.



The subtree that hosts the coordinator in (95a,b) is generated through subordination. Its formation must therefore involve discharge of a selectional requirement. We propose that the coordinator is a functional head that selects the XP it attaches to. Like other functional heads, a coordinator is a functor. Indeed, coordinators are *total* functors: they allow projection of *all* such properties. (Note that full transparency for projection explains why coordinators can be attached to nonmaximal categories.)

This puts coordinators in a large class of well-studied elements known as *linkers* (Den Dikken 2006, Philip 2012): functional heads that mark an independently existing relationship between two categories. An example is Mandarin *de* (Paul 2012):

(96) Mandarin

a. Meili$_{DP}$ **de** pengyou
Mary LNK friend
'Mary's friend'

b. benlai$_{AP}$ **de** yisi
original LNK meaning
'the original meaning'

c. [$_{S}$ ni jilai] **de** xin
you send LNK letter
'the letter you sent'

d. [$_{PP}$ dui wenti] **de** kanfa
towards problem LNK opinion
'an opinion about the problem'

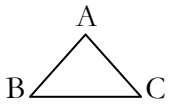
In asymmetric structures linkers are attached to the subordinated category and linearized between that category and the head of the larger structure. Thus, a linker connects two categories by being attached to one (XP in (97)) and pointing to the other (N in (97)) (Dik 1983, 1997, Philip 2012, 2013). 'Pointing to N', means 'appear on the same side of XP as N', not 'be adjacent to N'.

(97) a. [[XP LNK] N] c. *[[LNK XP] N]
b. [N [LNK XP]] d. *[N [XP LNK]]

Several authors have argued that coordinators should be analyzed as linkers (Dik 1983, Zwart 2009 and Philip 2012). Coordinators, like subordinating linkers, mark a relationship by linear intervention: they invariably intervene between two conjuncts (Dik 1997, Johannessen 1998 and Zwart 2009).

In asymmetric structures, a linker marks a bivalent relation. Coordinate structures are different. *A, B and C* expresses a *three-way* relation. This is true of the syntax, as on our account A, B and C are adjoined to each other. It is also true of the semantics, as the coordination as a whole denotes a set {A, B, C}. The three-way relation between A, B and C can be decomposed into three two-way relations:

(98)



Like other linkers coordinators mark relationships by pointing. Hence, in *A, B and-C* the coordinator marks two relations (A&C and B&C in (99)). A leftward shift of the coordinator *reduces* the number of relations it marks. In **A and B, C* only A&B is marked, while in **and A B C* no two-way relation is marked at all.

(99)

	A B &-C	*A &-B C	*&-A B C	A &-B &-C
A&B	*implied*	**marked**	unmarked	**marked**
A&C	**marked**	unmarked	unmarked	**marked**
B&C	**marked**	unmarked	unmarked	**marked**

Note that "co-project" and "co-members of set S" are transitive relationships. Therefore marking A&C and B&C implicitly marks A&B as well. It is of course possible to explicitly mark A&B as well, but this requires an additional coordinator, as in *A and B and C* (see (99)).

The distribution of coordinators in English follows from three constraints that interact in optimality-theoretic fashion (Prince & Smolensky 2004):

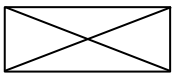
(100) a. No coordinators! (NoCo)
b. Mark coordinate relationships! (MkCo)
c. Explicitly mark coordinate relationships! (ExMk)

MkCo must universally dominates ExMk (the opposite ranking would obliterate the empirical effects of MkCo). We assume that in English NoCo is not ranked with respect to the other two constraints, leading to the the following three rankings:

(101) a. NoCo ≫ MkCo ≫ ExMk
b. MkCo ≫ NoCo ≫ ExMk
c. MkCo ≫ ExMk ≫ NoCo

We will consider the effects of these rankings for a coordinate structure with four members:

(102) A B
C D



On the ranking in (101a), it is more important to avoid coordinators than it is to mark any coordinate relations. Consequently, the coordination will be realized asyndetically, as *A, B, C, D*.

On the ranking in (101b), coordinators must be attached up to the point that all coordinate relations are marked (explicitly or implicitly). If a single coordinator is used, MkCo is satisfied if the coordinator is attached to the final conjunct, but not if it is attached to any preceding conjunct:

(103)

	A B C &-D	*A B &-C D	*A &-B C D	*&-A B C D
A&B	*implied*	*implied*	**marked**	unmarked
A&C	*implied*	**marked**	unmarked	unmarked
A&D	**marked**	unmarked	unmarked	unmarked
B&C	*implied*	**marked**	unmarked	unmarked
B&D	**marked**	unmarked	unmarked	unmarked
C&D	**marked**	unmarked	unmarked	unmarked

Attachment of more coordinators will not give a better performance on MkCo but does yield more violations of NoCo and is hence blocked on this ranking.

(104)

	*A B &-C &-D	*A &-B C &-D	*&-A B C &-D
A&B	*implied*	**marked**	*implied*
A&C	**marked**	*implied*	*implied*
A&D	**marked**	**marked**	**marked**
B&C	**marked**	*implied*	*implied*
B&D	**marked**	**marked**	**marked**
C&D	**marked**	**marked**	**marked**

(104)

	*A &-B &-C D	*&-A B &-C D	*&-A &-B C D
A&B	**marked**	*implied*	**marked**
A&C	**marked**	**marked**	unmarked
A&D	unmarked	unmarked	unmarked
B&C	**marked**	**marked**	unmarked
B&D	unmarked	unmarked	unmarked
C&D	unmarked	unmarked	unmarked

On the ranking in (101c), coordinators will be attached up to the point that all coordinate relationships are marked explicitly. This requires a three coordinators: ExMk is satisfied if a coordinator is attached to all non-initial conjuncts, as in the first column in (105).

(105)

	A &-B &-C &-D	*&-A B & C & D	*&-A &-B C &-D	*&-A &-B &-C D
A&B	**marked**	*implied*	**marked**	**marked**
A&C	**marked**	**marked**	*implied*	**marked**
A&D	**marked**	**marked**	**marked**	unmarked
B&C	**marked**	**marked**	*implied*	**marked**
B&D	**marked**	**marked**	**marked**	unmarked
C&D	**marked**	**marked**	**marked**	unmarked

Even though NoCo is ranked lowest, it still imposes constraints. In particular, it blocks attachment of *four* coordinators, as in *&-A &-B &-C &-D

Thus, three patterns are permitted: asyndetic coordination, attachment of the coordinator to the final conjunct, and attachment of coordinators to all non-initial conjuncts. Other patterns are correctly ruled out.

One issue remains: why should it be that asyndetic coordination triggers a sense of incompleteness or open-endedness? I speculate that this is due to the coordinate relation remaining unmarked. Marking a coordinate relation comes with the premise that the full set of elements of which the relation holds is marked. If so, the logic of interpretive competition predicts that lack of marking should have the effect that the relevant set may have additional members.

3.6 Conclusion

- Most syntactic structures are binary branching (see Kayne 1984).
- This follows from the Generalized Licensing Criterion, which requires (i) that subordination is licensed through the discharge of a selectional requirement and (ii) that no node created by subordination may be the locus of satisfaction of two or more selectional requirements.
- The Generalized Licensing Criterion predicts an exception to the binary branching constraint: it allows symmetric structures to be *n*-ary branching.
- We have argued that this exception indeed exists. Coordinate structures are arguably symmetric. They can indeed be *n*-ary branching.
- Coordinators are linkers that mark coordinate relationships.

4. Extraction asymmetries in Type A Coordination

The theory of coordination developed previously explains why the coordinate structure constraint should hold. However, various counterexamples to the coordinate structure constraint have been identified. Here, we consider one type of counterexample and show that it is only apparent: what looks like a coordinate structure is really a regular adjunction structure.

This section is based on Neeleman, A., and M. Tanaka (2024). Extraction Asymmetries Show that Type A Coordination is Adjunction. *Language* 100: 1–39.

4.1 Introduction

The theory of coordination developed in the previous handout captures the element condition as symmetry between conjuncts requires identical sets of licensing selectional requirements.

(1) In a coordinate structure, each conjunct must have the same set of unsatisfied licensing selectional requirements as its mother.

(2) *Coordinate Structure Constraint* (Ross 1967, Grosu 1973, De Vries 2017)
   a. The Conjunct Constraint: Conjuncts cannot be moved.
   b. The Element Constraint: Movement out of a coordinate structure is possible only if the moved category binds a trace in each conjunct.

(3) a. The madrigals which$_1$ Henry [[sings t_1] and [listens to t_1]] are mostly Venetian.
   b. *The lute which$_1$ Henry [[plays t_1] and [sings madrigals]] is warped.
   c. *The madrigals which$_1$ Henry [[plays the lute] and [sings t_1]] sound lousy.

Ross (1967: 168) already observed that there are apparent counterexamples to the CSC:

(4) Here's the whisky which$_1$ I [[went to the store] and [bought t_1]].

Ross further points out that non-ATB extraction requires a specific asymmetric interpretation of the coordinate structure (see also Lakoff 1986). Here, the first conjunct must describe an event that facilitates the event described by the second conjunct.

(5) a. *Here's the whisky which$_1$ I [[bought t_1] and [went to the store]].
   b. The madrigals which$_1$ Henry [[listens to t1] and [sings t_1]] are mostly Venetian.

Interpretive accounts assume that examples like 4 involve standard coordination and that non-ATB extraction is possible under specific interpretive conditions (Lakoff 1986, Deane 1991, Na & Huck 1992, Kehler 2002, Chaves 2012, Kubota & Lee 2015). *Syntactic accounts* assume that examples like (4) have a syntactic structure distinct from standard coordination, a structure that permits non-ATB extraction (Ross 1967, De Vos 2009, Weisser 2015, Bošković 2020).

Ross argued that in (4) the second conjunct starts out as a rationale clause. Given that English permits extraction from rationale clauses, as shown in (6), the acceptability of (4) follows.

(6) Here's the whisky which$_1$ I [[went to the store] [to buy t_1]].

This proposal can be updated by saying that the second part of an asymmetric coordination is not a *conjunct*, but an *adjunct* (see Brown 2017).

Extraction asymmetries show that the apparent right conjunct in examples like (4) is indeed an adjunct. The argument is based on eight experiments. Three experiments (IA, IIA and IVA) and a follow-up study (III) explore extraction from asymmetric coordination. Three additional experiments explore extraction from structures containing rationale clauses (IB, IIB, and IVB). A further experiment shows that the observed asymmetries cannot be ascribed to length of movement (V).

We leave type B and type C coordination for future research. We also put aside instances of contiguous coordination (e.g. *he sat and drank whisky*; Carden & Pesetsky 1977, De Vos 2005, 2007, Wiklund 2007, Brown 2017).

(7) A coordinate structure allows non-ATB extraction iff it denotes a sequence of events which fits normal conventionalized expectations (TYPE A), runs counter to conventionalized expectations (TYPE B), or is causative in nature (TYPE C).

(8) a. [How much]$_1$ can you [[drink t_1] and [not end up hung-over]]?
b. [Which shoes]$_1$ did Terry [[run in t_1] and [hurt her knee]]?

4.2 Background and predictions

4.2.1 Three arguments against the adjunction account

Schmerling (1972) showed, based on contrasts like 9, that type A coordinations cannot be derived from structures containing rationale clauses. Note, however, that this does not show that the apparent second conjunct in a type A coordination is not an adjunct.

(9) a. I went to the store to buy some whiskey, but I bought Ripple instead.
b. *I went to the store and bought some whiskey, but I bought Ripple instead.

Kehler (2002) argued that asymmetric coordination does not satisfy a basic test for subordination. Unlike subordinate clauses, the second conjunct in an asymmetric coordinate structure cannot be fronted:

(10) a. John went to the store and bought some whisky.
b. *And bought some whisky, John went to the store.

The argument is not conclusive, as there are other adjuncts that resist fronting. Absent a theory of contrasts like those in (12) and (13), it cannot be decided whether (10b) violates the CSC or involves an unfrontable adjunct.

(11) a. We will provide you with a laptop, so (that) you can make the most of your time here.
b. So (that) you can make the most of your time here, we will provide you with a laptop.

(12) a. You never collected your new laptop, so (that) you couldn't do your work.
b. *So (that) you couldn't do your work, you never collected your new laptop.

(13) a. This analysis is problematic, in that it relies on numerous stipulations.
b. *In that it relies on numerous stipulations, this analysis is problematic.

Lakoff (1986) noted that there can be more than two conjuncts in asymmetric coordinate structures.

(14) This is [the cake]$_1$ that Harry [went to the store], [bought t_1], [loaded $t_1$in his car], [came home], and [put t_1 in the fridge].

In (14), extraction is from the second, third and fifth conjuncts. Lakoff suggests that this is problematic on the standard version of the CSC and a syntactic approach to asymmetric coordination. It is neither compatible with an analysis of (14) in which the five conjuncts form a single coordinate structure, nor with an analysis in which the four noninitial conjuncts are treated as adjuncts. However, Weisser (2015) argues that (14) can be understood if analyzed as a coordination with three terms:

(15) This is [the cake]$_1$ that Harry
[[went to the store], [bought t_1]
[loaded t_1 in his car]
[came home] and [put t_1 in the fridge].]

(16) This is [the cake]$_1$ that Harry
[[went to the store] [to buy t_1]
[loaded t_1 in his car]
and [came home] [to put t_1 in the fridge].]

CSC violations are predicted if any of the gaps in (14) is filled. This effect is indeed found (judgments based on a questionnaire study, n=16; see also Kehler 2002 and Bošković 2020):

(17) a. This is the cake that Harry [went to the store, bought *t*], [loaded *t* in his car], [came home and put *t* in the fridge]. (3.44 out of 5)
b. *This is the cake that Harry [went to the store, made his purchase], [loaded *t* in his car], [came home and put *t* in the fridge]. (1.81)
c. ??This is the cake that Harry [went to the store, bought *t*], [loaded the box in his car], [came home and put *t* in the fridge]. (2.13)

d. *This is the cake that Harry [went to the store, bought *t*], [loaded *t* in his car], [came home and put the box in the fridge]. (1.81)

(18) a. This is the cake that Harry [went to the store to buy *t*], [loaded *t* in his car], and [came home to put *t* in the fridge]. (4.4)
b. *This is the cake that Harry [went to the store to make his purchase], [loaded *t* in his car], and [came home to put *t* in the fridge]. (2.12)
c. ??This is the cake that Harry [went to the store to buy *t*], [loaded the box in his car], and [came home to put *t* in the fridge]. (2.68)
d. ??This is the cake that Harry [went to the store to buy *t*], [loaded *t* in his car], and [came home to put the box in the fridge]. (2.56)

Thus, Lakoff's (1986) argument from examples like (14) is less strong than it appears. Note, though, that (14) contains only one coordinator, while one would expect two additional coordinators. We cannot address this issue here and leave it as a potential problem for the syntactic account.

(19) This is [the cake]$_1$ that Harry [[went to the store] and [bought t_1]], [loaded $t_1$in his car], and [[came home] and [put t_1 in the fridge]].

4.2.2 Extraction asymmetries

Movement is subject to island constraints. As an example, compare *wh*-extraction from the subject of a finite clause with *wh*-extraction from the object:

(20) a. Who did you meet [a friend of t_{WH}] at the airport?
b. *Who did [a friend of t_{WH}] meet you at the airport?

The classical view is that islands are syntactic (Ross 1967; Chomsky 1973, 1977, 1986, 1995). Alternative approaches treat islands as semantic (Erteschik-Shir 1973, Kuno 1987, Goldberg 2006, Ambridge et al. 2014) or processing phenomena (Deane 1991, Kluender 1991, Hofmeister & Sag 2010). This debate is not immediately relevant here. For concreteness' sake, we will assume the classical view.

The literature identifies several factors that determine acceptability of extraction from adjuncts. The first is height of attachment (VP-internal/external; Huang 1982, Chomsky 1986, Truswell 2011).

(21) a. They$_1$ were too angry [PRO$_{1/2}$ to hold the meeting].
b. [Which meeting] were they too angry [PRO$_{1/*2}$ to hold t_{wh}]?

(22) a. We should [meet the students in Caroline's room].
b. [In whose room]$_1$ should we meet the students t_1?
c. ?[Whose room]$_1$ should we meet the students [in t_1]?

(23) a. We should [meet the students] in Caroline's opinion.
b. [In whose opinion]$_1$ should we meet the students t_1?
c. *[Whose opinion]$_1$ should we meet the students [in t_1]?

A second factor concerns the displaced category. While complements allow extraction of arguments and adverbials, adjuncts only marginally allow extraction of arguments and block extraction of adverbials altogether (Lasnik and Saito 1984, 1992; Chomsky 1986; Cinque 1990; Borgonovo 1994; Postal 1998; Szabolcsi 2006; Truswell 2011). Thus, (VP-internal) adjuncts are *selective* islands (see Szabolcsi 2006 and Abrusán 2014 for discussion and references).

(24) a. She wanted [to measure the distance between A and B with a high level of accuracy].
b. [What distance]$_1$ did she want [to measure t_1 with a high level of accuracy]?
c. [With what level of accuracy]$_1$ did she want [to measure the distance between A and B t_1]?

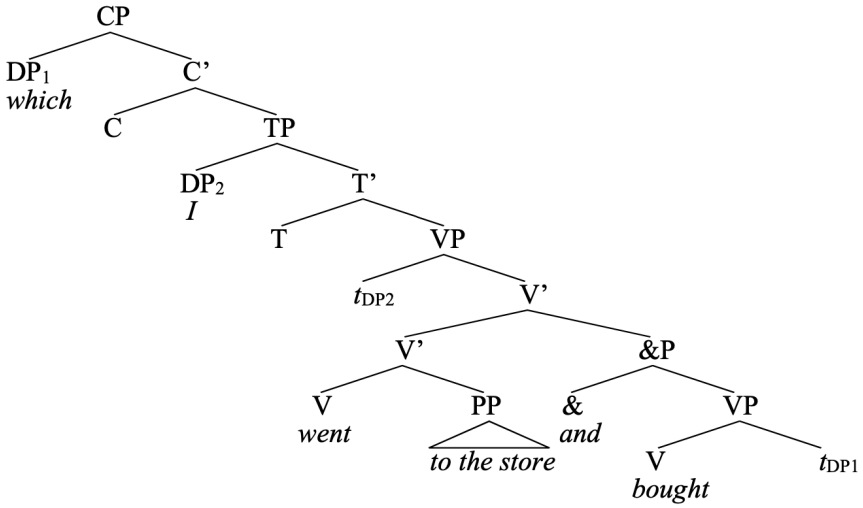
(25) a. She went out [to measure the distance between A and B with a high level of accuracy].
b. ?[What distance]$_1$ did she go out [to measure t_1 with a high level of accuracy]?
c. *[With what level of accuracy]$_1$ did she go out [to measure the distance between A and B t_1]?

A third factor is the nature of the verb that heads the host structure. We will explore this later.

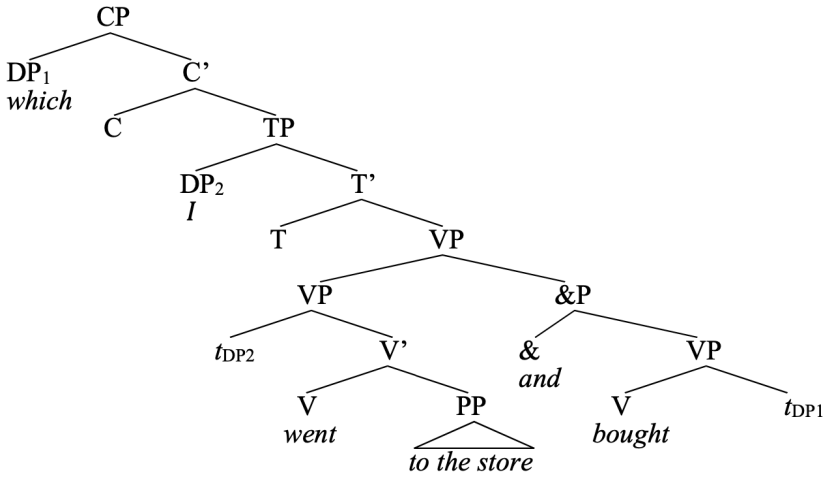
4.2.3 The adjunction account and its predictions

We claim that the second conjunct in an example like (4) is a VP-internal adjunct. An open question is what counts as VP-internal. We assume that categories adjoined to VP are VP-external (May 1985, Chomsky 1986), and so transparent adjuncts are attached lower, as in 26a. Alternatively, attachment *to* VP (as in 26b) could be deemed low enough for transparency (Altshuler & Truswell 2022).

(26) a.



b.



Given that the apparent *right* conjunct is analyzed as an adjunct, it should marginally permit extraction of arguments and to block extraction of modifiers. Given that the apparent *left* conjunct is analyzed as the main predicate, it should freely permit extraction of both arguments and modifiers. Thus, we predict an interaction between the *extraction domain* (left vs. right conjunct) and the *extracted category* (argument vs. modifier). N.B.: many existing analyses of type A coordination allow extraction from the *right* conjunct but rule out or render marked extraction from the *left* conjunct (Deane 1991; Na & Huck 1992; Kehler 2002; De Vos 2005, 2009; Weisser 2015; Bošković 2020).

A third prediction concerns the predicate that heads the left conjunct. On the adjunction analysis, the right conjunct must be classified as subject-oriented. This is relevant because it has been claimed that extraction from subject-oriented depictives improves if the matrix verb is unaccusative, as opposed to transitive (Borgonovo 1994, Cormack & Breheny 1994, Borgonovo & Neeleman 2000, Fábregas & Jiménez-Fernández 2016).

(27) a. John$_1$ killed Bill [thinking about Mary]$_1$.
b. *Who did John$_1$ kill Bill [thinking about]$_1$?
c. Bill$_1$ died t$_1$ [thinking about Mary]$_1$.
d. ?Who did Bill$_1$ die t$_1$ [thinking about]$_1$?

Thus, it is predicted that the nature of the predicate in the left conjunct will affect the acceptability of extraction from the right conjunct to the same extent as the nature of the main predicate affects extraction from a subject-oriented adjunct.

There are good reasons to investigate the influence of the predicate. Type A coordination is felicitous with a motion verb accompanied by a directional modifier in the left conjunct (Ross 1967, Schmerling 1972, Lakoff 1986, Deane 1991, De Vos 2005, Weisser 2015). Such predicates are unaccusative (Hoekstra 1984, Levin and Rappaport Hovav 1995). It is controversial, however, whether the verb in the left conjunct can be transitive (Schmerling 1972, Weisser 2015 vs. De Vos 2005, Brown 2017)

If the predicate effect is real, there is a question of what explains it. We discuss this later.

4.3 The predicate effect and the extraction domain effect

4.3.1 Information relevant to all experiments

In what follows, we report on eight experiments that suggest that type A coordination is adjunction.

- For each experiment, we recruited eighty self-reporting native speakers of English through *Prolific*.
- All experiments involved acceptability judgement tasks created and conducted using *Gorilla*. Participants were asked to indicate the acceptability of items on a seven-point Likert scale.
- We transformed raw acceptability scores into z-scores and ran a linear mixed-effects model on the resulting data, using R's lme4 package. Models included random intercepts for both subject and item. We also included by-subject random slopes for fixed effects and their interaction, unless the model failed to converge, in which case we removed the by-subject random slope for the interaction.
- The lmerTest package was used to calculate *t*-statistics and *p*-values based on Satterthwaite's approximation. Post-hoc pairwise comparisons were conducted using the emmeans package.
- Six of our experiments were coupled: each item in the experiment exploring extraction from type A coordination formed a minimal pair with its counterpart in the experiment on extraction from structures containing rationale clauses. We ran a correlation analysis on the results of these coupled experiments to determine whether the acceptability of extraction from a structure containing a rationale clause was a good predictor of the acceptability of extraction from type A coordination.

4.3.2 Experiments IA and IB

We claim that in type A coordination extraction from the second conjunct is possible because this conjunct is a low adjunct. Exp. IA tests two predictions that follow from this.

- If the right conjunct is an adjunct, it is a subject-oriented adjunct. Extraction from such adjuncts is facilitated by the matrix verb being unaccusative. Thus, argument extraction from the right conjunct should be sensitive to the nature of the verb in the left conjunct (PREDICTION 1).
- If the right conjunct is an adjunct, the left conjunct is the matrix VP. A matrix VP is transparent for extraction, while low adjuncts are selective islands. Thus, argument extraction from the left conjunct should be better than argument extraction from the right conjunct (PREDICTION 2).

These predictions presuppose that the predicate effect and the extraction domain effect are real in uncontroversial adjunction structures. Exp. IB therefore tests parallel predictions pertaining to structures containing (subject-oriented) rationale clauses.

- Argument extraction from a (subject-oriented) rationale clause should be sensitive to the nature of the matrix verb (PREDICTION 1').
- Argument extraction from the matrix VP should be better than argument extraction from a rationale clause (PREDICTION 2').
- There should be a highly positive correlation between the results of experiments IA and IB (PREDICTION 3).

Exp. IA had a 2×2 factorial design in which we manipulated two factors: VERB CLASS (motion vs. transitive verb) and GAP POSITION (left vs. right conjunct). Exp. IB had a similar design, with VERB CLASS (motion vs. transitive verb) and GAP POSITION (matrix clause vs. adjunct) as the factors manipulated. In the motion verb condition, we used predicates that express directed motion. In the transitive verb condition, we used verbs with an object understood as instrumental for the event described by the second conjunct (in exp. IA) and the rationale clause (in exp. IB).

A set of sample test items is given below (L/R=extraction from Left/Right conjunct; M/T=Motion/Transitive predicate; Mx/Adj=extraction from Matrix VP/Adjunct VP). The bracketed sentences were given as context to make sure that participants construed the test items as intended.

(28) [Ali travelled to Kinshasa in order to fight the heavyweight champion.]
  a. What city did Ali travel to and fight the heavyweight champion? (L | M)
  b. What boxer did Ali travel to Kinshasa and fight? (R | M)

(29) [Ali put on his favourite gloves in order to fight the heavyweight champion.]
  a. What gloves did Ali put on and fight the heavyweight champion? (L | T)
  b. What boxer did Ali put on his favourite gloves and fight? (R | T)

(30) [Ali travelled to Kinshasa in order to fight the heavyweight champion.]
  a. What city did Ali travel to in order to fight the heavyweight champion? (Mx | M)
  b. What boxer did Ali travel to Kinshasa in order to fight? (Adj | M)

(31) [Ali put on his favourite gloves in order to fight the heavyweight champion.]
  a. What gloves did Ali put on in order to fight the heavyweight champion? (Mx | T)
  b. What boxer did Ali put on his favourite gloves in order to fight? (Adj | T)

For each experiment, we created twelve sets of test items, and so there were forty-eight test items in total (twelve sets × four conditions). The experiments had a Latin square design with four lists that contained four practice items, twelve test items, and twenty-four fillers. Non-test items were pseudorandomized per participant, with the test items interspersed with fillers.

A linear mixed effect model was fitted to the results of exp. IA, with VERB CLASS, GAP POSITION, and their interaction as fixed effect predictors. It revealed a significant main effect of VERB CLASS (Est.=0.80, SE=0.06, t=12.64, p<.001), reflecting the fact that motion verbs facilitated extraction from both the right conjunct (Est.=0.35, SE=0.06, t=5.48, p<.001) and the left conjunct (Est.=0.80, SE=0.06, t=12.64, p<.001).

We also observed a significant main effect of GAP POSITION (Est.=–0.61, SE=0.06, t=–9.15, p<.001): for both verb classes, extraction from the right conjunct was rated lower. Pair-wise comparisons show the effect to be stronger in the motion verb condition (Est.=0.61, SE=0.06, t=9.15, p<.001) than with transitive verbs (Est.=0.15, SE=0.06, t=2.32, p<.05), resulting in a significant interaction between GAP POSITION and VERB CLASS (Est.=0.45, SE=0.09, t=5.07, p<.001).

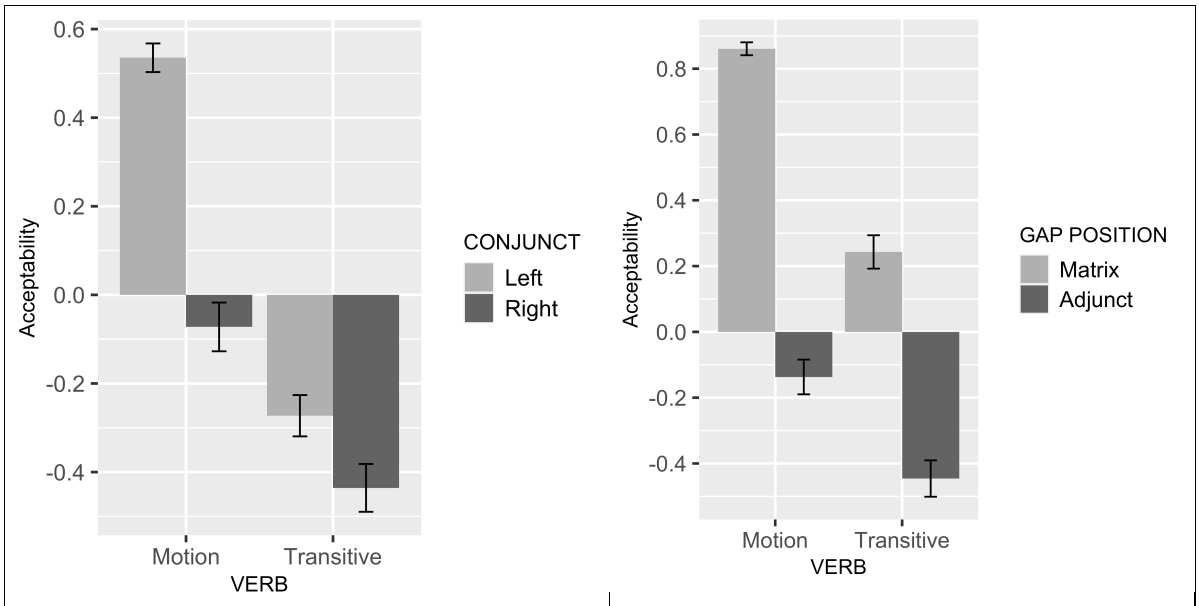


Figure 1: Mean acceptability by condition (z-scores; GAP POSITION × VERB CLASS) in exp. IA

Figure 2: Mean acceptability by condition (z-scores; GAP POSITION × VERB CLASS) in exp. IB

A linear mixed effect model with VERB CLASS, GAP POSITION, and their interaction as fixed effect predictors was also fitted to the results of exp. IB, showing significant main effects of VERB CLASS (Est.=0.62,

SE=0.07, t=8.56, p<.001) and of GAP POSITION (Est.=–0.99, SE=0.06, t=–14.41, p<.001), and a significant interaction between VERB CLASS and GAP POSITION (Est.=0.30, SE=0.09, t=3.119, p<.01).

The VERB CLASS effect is due to motion verbs aiding extraction from both the rationale clause (Est.=0.31, SE=0.06, t=4.87, p<.001) and the matrix clause (Est.=0.62, SE=0.07, t=8.56, p<.001). The effect of GAP POSITION shows that for both verb classes extraction from the rationale clause scored lower than extraction from the matrix clause. Pairwise comparison revealed this contrast to be larger with motion verbs (Est.=0.99, SE=0.06, t=14.41, p<.001) than with transitive verbs (Est.=0.68, SE=0.08, t=8.25, p<.001).

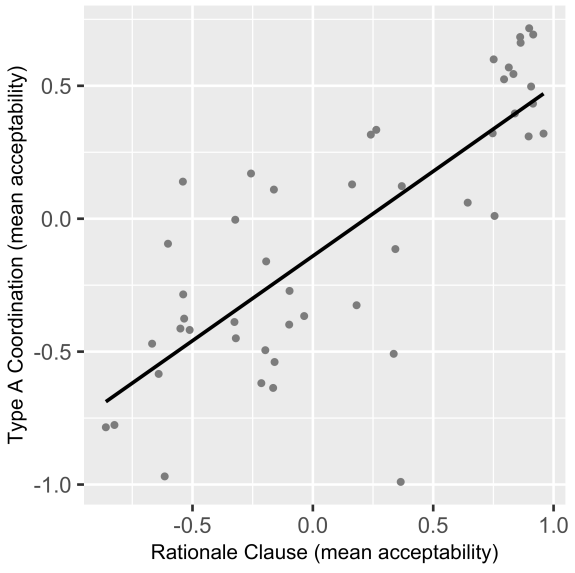


Figure 3: Correlation between the acceptability of items in exp. IB and related items in exp. IA

This resulted in a significant interaction between the two factors. However, the effect was smaller than in exp. IA. (The difference in exp. IB between extraction from a matrix clause and a rationale clause was much sharper in the transitive verb condition than the difference in exp. IA between extraction from the left and the right conjuncts.)

Given that experiments IA and IB yielded similar results, it is not surprising that acceptability of extraction from type A coordinate structures in exp. IA and from their rationale clause counterparts in exp. IB show a highly positive correlation (tau=0.5549645). About 60 % of the variance in the acceptability of type A coordination items can be accounted for based on the acceptability of related rationale clause items (r^2=0.6041). This positive correlation is highly significant (t=8.378, df=46, p<.001).

4.3.3 Discussion

BASIC PREDICTIONS. The hypothesis that the right conjunct in type A coordination is an adjunct predicts a parallel between extraction from type A coordinate structures and extraction from rationale clauses. Experiments IA and IB confirm this parallel.

In type A coordination, argument extraction from the right conjunct should be sensitive to the nature of the verb in the left conjunct (prediction 1). Similarly, argument extraction from a rationale clause should be sensitive to the nature of the matrix verb (prediction 1'). Both predictions are borne out.

In type A coordination argument extraction from the left conjunct should be better than argument extraction from the right conjunct (prediction 2). In the same vein, argument extraction from the matrix VP should be better than argument extraction from a rationale clause (prediction 2'). Again, both predictions are borne out.

Finally, there should be a strong positive correlation between the results of experiments IA and IB (prediction 3). A highly significant positive correlation was indeed observed.

THEORETICAL INTERPRETATION. Why should the predicate effects exist?

Subject-oriented secondary predicates are attached higher than object-oriented ones (Williams 1980, Rothstein 1983, Demonte 1987, McNulty 1988, Nakajima 1990, Carrier & Randall 1992, Bowers 2001, Janke & Neeleman 2012).

(32) a. John$_1$ sketched [the model]$_2$ [nude]$_2$ [drunk as a skunk]$_2$.
b. John$_1$ sketched [the model]$_2$ [nude]$_2$ [drunk as a skunk]$_1$.
c. John$_1$ sketched [the model]$_2$ [nude]$_1$ [drunk as a skunk]$_1$.
d. *John$_1$ sketched [the model]$_2$ [nude]$_1$ [drunk as a skunk]$_2$.

We assume that while object-oriented depictives are attached VP-internally (see 33a), subject-oriented depictives are (typically) adjoined to VP (see 33b).

(33) a. DP_1 V [$_{VP}$ DP_2 [t_V XP_2]] (transitive construction; object-oriented predicate)
b. DP_1 [$_{VP}$ [$_{VP}$ V DP_2] XP_1] (transitive construction; subject-oriented depictive)

Explanation: (i) The c-command constraint on predication implies that object-oriented depictives must be attached VP-internally, while subject-oriented depictives can be attached either to VP or VP-internally. (ii) A locality constraint on predication (*c-subjacency*; Wiliams 1980) then militates against low attachment of subject-oriented predicates (c-subjacency is violated if a maximal projection dominates a secondary predicate but not the associated argument).

C-subjacency expresses a preference, rather than an absolute requirement, so that subject-oriented secondary predicates can be attached VP-internally, but at the cost of reduced acceptability. Prediction 1: Extraction from subject-oriented secondary predicates should be worse than extraction from object-oriented secondary predicates, as the latter only allow extraction in their dispreferred attachment site. This has been argued to be correct in Brown 2017. Prediction 2: extraction from subject-oriented secondary predicates requires VP-internal attachment. One way to test this is with *do so* substitution:

(34) [*Anthony*: John drove Mary crazy talking about etiquette, and Bill did so talking about something similarly trivial.]
a. *Bernadette*: Oh. (?)What did Bill drive Mary crazy talking about?
b. *Bernadette*: Oh. *What did Bill do so talking about?

In unaccusative and reflexive constructions, it is possible for an apparently subject-oriented secondary predicate to be attached VP-internally without violating c-subjacency:

(35) a. DP_1 V [$_{VP}$ t_1 [t_V XP_1]] (unaccusative construction, VP-internal predicate)
b. DP_1 V [$_{VP}$ $REFL_1$ [t_V XP_1]] (reflexive construction, VP-internal predicate)

This captures the predicate effect. The rationale clauses in exp. IA and the apparent right conjuncts in exp. IB are subject-oriented. When the predicate is transitive, VP-internal attachment (necessary for extraction) induces a c-subjacency violation. When the predicate is a motion verb, the rationale clauses in exp. IA and the apparent right conjuncts in exp. IB can be attached VP-internally without penalty.

AN UNEXPECTED PREDICATE EFFECT. Our results showed a broader predicate effect than expected. In experiments IA and IB, the unaccusativity of the predicate in the left domain affected extraction not only from the right domain, but also from the left domain itself.

The effect may partly result from a systematic contrast in extraction sites. In both experiments, extraction from the left domain targeted an object when the predicate was transitive and the complement of a preposition when the predicate was unaccusative.

This cannot be the whole story, though. The improvement observed in type A coordination was much larger than that observed in structures with rationale clauses (compare figures 1 and 2). Why? Our conjecture is that type A coordination competes in parsing with regular coordination, while there is no comparable competing parse in structures with rationale clauses.

The CSC implies that a deactivated filler must be reactivated when the parser encounters a coordinator (Wagers & Phillips 2009 and Kim et al. 2020). Type A coordinate structures are likely to be analyzed as regular coordinations initially. If so, the sequence of operations just described will be triggered when an element is extracted from the initial VP, leading to backtracking if it turns out that there is no insertion site for a trace in the right conjunct.

Misanalyzed ambiguities lead to repair whose impact depends on the ease of diagnosis (Fodor & Inoue 1994; see Fujita 2021 for recent discussion). We speculate that diagnosis is easier if there is a nonisomorphic alternative to the incorrect initial parse. Adjunction of a constituent *to* VP is isomorphic to coordination of that constituent with VP, but attachment *within* VP is not.

It is here that the nature of the predicate in the first conjunct becomes relevant. Only if that predicate is unaccusative, a straightforward alternative attachment site is available for the second conjunct, namely as a VP-internal adjunct (see 35a). If the verb in the first conjunct is transitive, VP-internal attachment of the second conjunct violates c-subjacency and is hence less accessible. A more serious backtracking effect ensues, leading to lower experimental scores.

4.4 More on the predicate effect

4.4.1 Experiments IIA and IIB

Like motion verbs, posture verbs accompanied by a directional particle (e.g. change-of-state *sit down*) allow an agentive construal while being unaccusative in their syntax (Levin & Rappaport Hovav 1995). Hence, in type A coordination, posture verbs should pattern with motion verbs, rather than with transitives, which inhibit extraction (see Lakoff 1986; Deane 1991; De Vos 2005, 2009; Weisser 2015).

We tested this in experiments IIA and IIB. In exp. IIA, VERB CLASS in the left conjunct of a type A coordination was a within-subject factor with three values (motion, posture, and transitive). Exp. IIB had a parallel design. We manipulated the class of verb in the main clause and tested the acceptability of argument extraction from a rationale clause. Each test item in exp. IIA constituted a minimal pair with its counterpart in exp. IIB.

(36) a. [John hurried to the airport in order to welcome the guest from Berlin.]
What guest did John hurry to the airport and welcome? (motion)
b. [John stood up in order to welcome the guest from Berlin.]
What guest did John stand up and welcome? (posture)
c. [John opened the car door in order to welcome the guest from Berlin.]
What guest did John open the car door and welcome? (transitive)

(37) a. [John hurried to the airport in order to welcome the guest from Berlin.]
What guest did John hurry to the airport in order to welcome? (motion)
b. [John stood up in order to welcome the guest from Berlin.]
What guest did John stand up in order to welcome? (posture)
c. [John opened the car door in order to welcome the guest from Berlin.]
What guest did John open the car door in order to welcome? (transitive)

In each experiment, there were thirty-six test items (12 sets × 3 conditions) distributed across three lists in a Latin Square fashion. Each list also contained four practice items and twenty-four fillers. Nonpractice items were pseudorandomized per participant. Predictions:

- In exp. IIA, motion and posture verbs (as compared to transitives) should facilitate extraction from the second conjunct in type A coordination (PREDICTION 4).
- In exp. IIB, motion and posture verbs (as compared to transitives) should facilitate extraction from rationale clauses (PREDICTION 4').
- Finally, there should be a highly positive correlation between extraction from adjuncts in exp. IIB and extraction from the second conjunct in type A coordination in exp. IIA (PREDICTION 5).

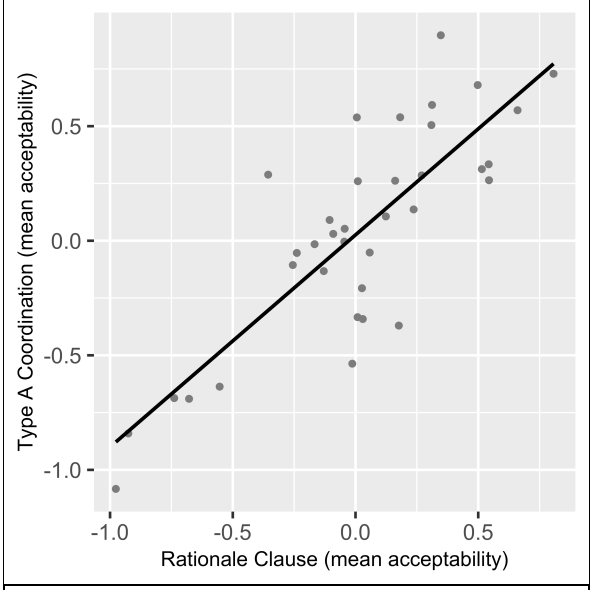


Figure 4: Correlation between extraction from rationale clauses (IIB) and Type A coordination (IIA)

The results were largely as expected. Mean acceptability per item in experiments IIB and IIA showed a strong positive correlation (r=0.821). About 67% of the variance in the acceptability of extraction from type A coordination can be accounted for based on the acceptability of extraction from rationale clauses (r^2=0.6652). This correlation is highly significant (t=8.399, df=34, p<.001).

This strong positive correlation would lead one to expect that extraction patterns from rationale clauses and type A coordination should mirror each other in detail. To check this, we consider the results of experiments IIA and IIB in more detail, starting with the former.

A linear mixed effects model was fitted to the results of exp. IIA with VERB CLASS as a fixed effect predictor. The model revealed that extraction from the right conjunct of a Type A coordination was easier when the left conjunct was headed by a motion

verb (Est.=0.32, SE=0.06, t=4.78, p<.001) or a posture verb (Est.=0.57, SE=0.06, t=8.80, p<.001), as opposed to a transitive verb.

A post-hoc Tukey test revealed significant differences in all pair-wise comparisons (p<.001), which indicates that while both motion and posture verbs facilitated extraction, there was a significant difference between them: items with posture verbs were on average rated higher than items with motion verbs (Est.=0.25, SE=0.05, t=4.33, p<.001).

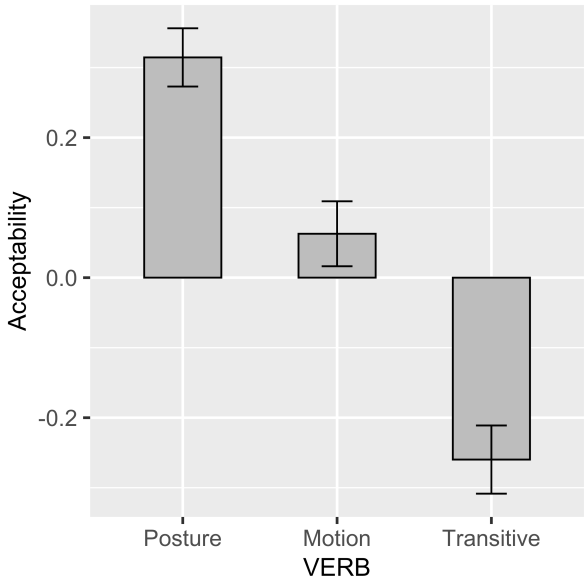


Figure 5: Mean acceptability by VERB CLASS (z-scores) in experiment IIA

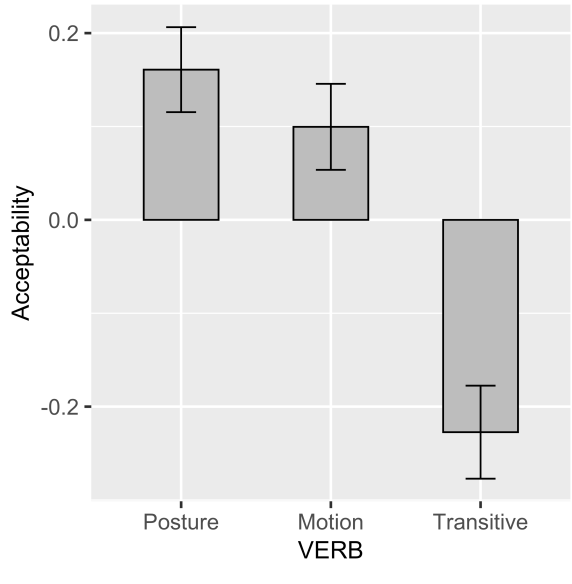


Figure 6: Mean acceptability by VERB CLASS (z-scores) in experiment IIB

A linear mixed effects model was also fitted to the results of exp. IIB with VERB CLASS as a fixed effect predictor. The model revealed that extraction from rationale clauses was easier when the matrix VP was headed by a motion verb (Est.=0.32, SE=0.07, t=4.35, p<.001) or a posture verb (Est.=0.38, SE=0.06, t=6.16, p<.001), rather than a regular transitive verb.

A post-hoc Tukey test displayed that there were significant differences between (i) motion verbs and nonmotion transitives (p<.001) and (ii) posture verbs and nonmotion transitives (p<.001). However, there was no significant difference between motion verbs and posture verbs (Est.=−0.05, SE=0.06, t=−0.86, p >.6).

These findings confirm all predictions tested. In both experiments, we find that extraction from the right-hand domain is easier when the verb in the left-hand domain is a motion or posture verb (predictions 4 and 4'). Thus, there is a highly positive correlation between extraction from rationale clauses and in type A coordination (prediction 5). But there is an unexpected wrinkle in the data: posture verbs had a greater impact on acceptability than motion verbs in type A coordination. Why?

Suggestion: participants may occasionally have analyzed type A coordination with posture verbs as contiguous coordination (e.g. *Amara sat and read a book*), which shows no island effects. In exp. IIA, posture verbs were often accompanied by just a particle (in line with a claim in De Vos 2005, 2009). Wiklund argues that a full directional PP must be included to rule out contiguous coordination.

4.4.2 Experiment III

To test whether the above suggestion explains the slight discordance between experiments IIA and IIB, we conducted a follow-up study to exp. IIA. As before, VERB CLASS was manipulated as a within-subject factor with three values (posture, motion, and nonmotion transitive). Most items were recycled from exp. IIA. However, posture verbs were accompanied by a directional PP, as in (38). Other than that, the experiment was identical to exp. IIA.

(38) [Sally sat down on her chair in order to sign the contract.]
What contract did Sally sit down on her chair and sign? (posture)

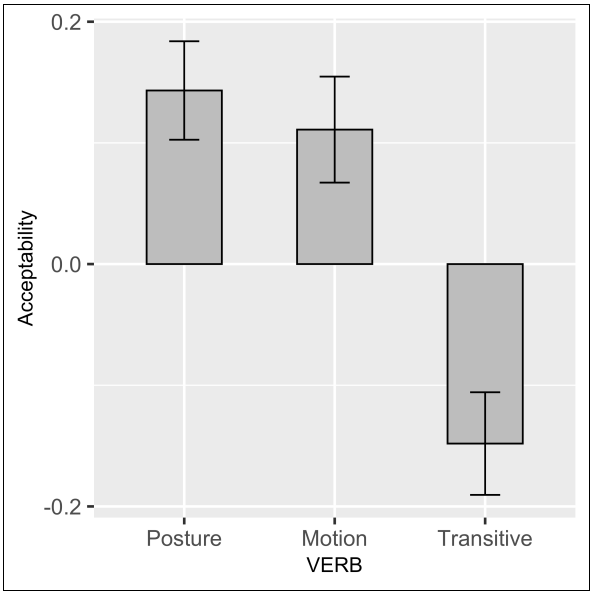


Figure 7: Mean acceptability by VERB CLASS (z- scores) in experiment III

A linear mixed effects model was fitted to the results of exp. III with VERB CLASS as a fixed effect predictor. It revealed a main effect of VERB CLASS. As shown in Figure 7, both posture verbs and motion verbs aided extraction from the right conjunct in Type A coordination (posture: Est.=0.29, SE=0.05, t=5.34, p<.001; motion: Est.=0.26, SE=0.05, t=4.65, p <.001).

A post-hoc Tukey test revealed that there were significant differences in acceptability between (i) the motion and transitive conditions (p<.001) and (ii) the posture and transitive conditions (p<.001). However, as predicted, there was no significant difference between the motion and posture conditions (Est.=−0.02, SE=0.05, t=−0.05, p >.8).

The results of exp. III thus match the results of exp. IIB. Extraction from the right conjunct in type A coordination and extraction from rationale clauses exhibit the same pattern: Posture ≈ Motion > Nonmotion Transitive, suggesting that the correlation between extraction from type A coordination and extraction from rationale clauses is even stronger in figure 4.

4.5. The extracted category effect

4.5.1 Experiments IVA and IVB

VP-internal adjuncts are selective islands: they impose a small penalty on argument extraction and block extraction of adverbials and other adjuncts. The hypothesis that type A coordination is adjunction thus makes two predictions (tested in exp. IVA):

- The right conjunct in type A coordination should marginally allow extraction of arguments, but not allow extraction of adjuncts (PREDICTION 6).
- The left conjunct in type A coordination should freely permit extraction of adjuncts as well as arguments (PREDICTION 7).

This behavior should be mirrored by that of structures containing rationale clauses (tested in exp. IVB):

- A rationale clause should marginally allow extraction of arguments, but not allow extraction of adjuncts at all (PREDICTION 6').
- The VP in which the rationale clause appears should freely permit extraction of adjuncts as well as arguments (PREDICTION 7').
- There should be a highly positive correlation between the results of experiments IVA and IVB (PREDICTION 8).

Type A coordination		
	Right Conj.	Left Conj.
Argument	?	✓
Adjunct	*	✓
	↑ Prediction 6	↑ Prediction 7

VP containing rationale clause		
	Adjunct	Matrix VP
Argument	?	✓
Adjunct	*	✓
	↑ Prediction 6'	↑ Prediction 7'

Table 1: Predictions of the adjunction analysis

Exp. IVA had a 2×2 factorial design in which we manipulated CONJUNCT (left vs. right) and EXTRACTED CATEGORY (argument vs. adjunct). Exp. IVB had a 2×2 factorial design in which we manipulated GAP POSITION (matrix clause vs. rationale clause) and EXTRACTED CATEGORY (argument vs. adjunct). Sample test items are given in 39 (for exp. IVA) and in 40 (for exp. IVB). Moved adjuncts had unambiguous launching sites. To ensure that examples with argument and adjunct extraction were equally complex we included an adjunct in conjuncts from which an argument was extracted.

(39) [Ali travelled to Kinshasa on a jumbo jet in order to defeat the heavyweight champion with a well-timed right hook.]

a. Ken knew the city which Ali travelled to on a jumbo jet and defeated the heavyweight champion. (Left | Argument)
b. Ken saw the jumbo jet on which Ali travelled to Kinshasa and defeated the heavyweight champion. (Left | Adjunct)
c. Ken knew the heavyweight champion who Ali travelled to Kinshasa and defeated with a well-timed right hook. (Right | Argument)
d. Ken noticed the well-timed right hook with which Ali travelled to Kinshasa and defeated the heavyweight champion. (Right | Adjunct)

(40) [Ali travelled to Kinshasa on a jumbo jet in order to defeat the heavyweight champion with a well-timed right hook.]

a. Ken knew the city which Ali travelled to on a jumbo jet in order to defeat the heavyweight champion. (Matrix | Argument)
b. Ken saw the jumbo jet on which Ali travelled to Kinshasa in order to defeat the heavyweight champion. (Matrix | Adjunct)
c. Ken knew the heavyweight champion who Ali travelled to Kinshasa in order to defeat with a well-timed right hook. (Rationale | Argument)
d. Ken noticed the well-timed right hook with which Ali travelled to Kinshasa in order to defeat the heavyweight champion. (Rationale | Adjunct)

For each experiment, we created twelve sets of four test items distributed across four lists along with four practice items and twenty-four fillers. Nonpractice items were pseudorandomized per participant.

A linear mixed effects model was fitted to the results of exp. IVA, with CONJUNCT, EXTRACTED CATEGORY and their interaction as fixed effect predictors. It revealed a main effect of CONJUNCT (Est.=0.86, SE=0.08, t=10.715, p <.001) – extraction from the left conjunct scored much higher than extraction from the right conjunct, with both adjuncts (Est.=0.86, SE=0.08, t=10.71, p<.001) and arguments (Est.=0.23, SE=0.07, t=2.92, p<.01).

EXTRACTED CATEGORY did not have a significant effect (Est.=−0.10, SE=0.06, t=−1.576, p>.1), but there was a highly significant interaction between CONJUNCT and EXTRACTED CATEGORY (Est.=0.62, SE=0.11, t=5.70, p <.001). In the right conjunct condition, adjunct extraction scored much lower than argument extraction (Est.=−0.52, SE=0.08, t=−6.06, p<.001). The left conjunct condition showed no significant difference (Est.=0.10, SE=0.06, t=1.57, p>.1).

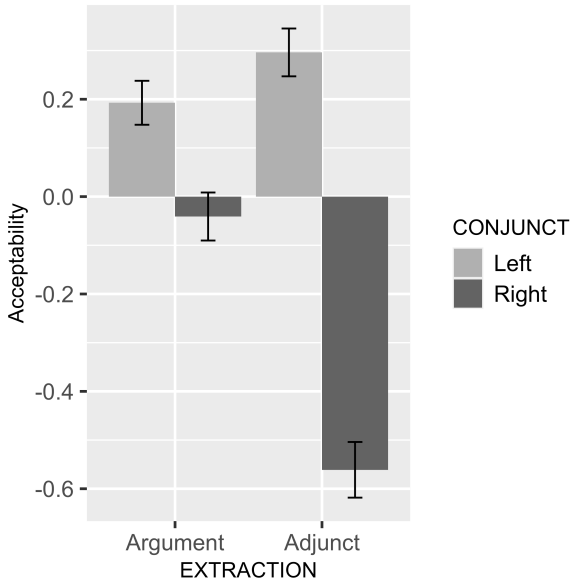


Figure 8: Mean acceptability by condition (z-scores; CONJ. × EXTR. CATEGORY) in exp. IVA

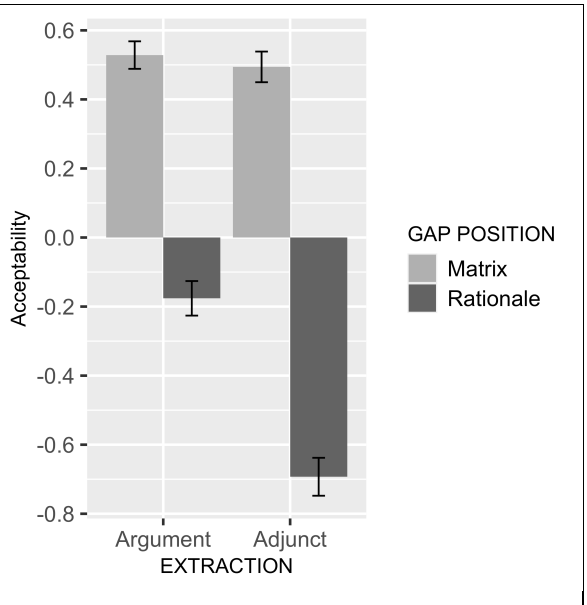


Figure 9: Mean acceptability by condition (z-scores; GAP POS. × EXTR. CATEGORY) in exp. IVB

A linear mixed effects model was also fitted to the results of exp. IVB, with GAP POSITION, EXTRACTED CATEGORY and their interaction as fixed effect predictors. It revealed a main effect of GAP POSITION (Est.=1.19, SE=0.07, t=15.506, p<.001) – extraction from the matrix clause scored much higher than extraction from a rationale clause both with adjuncts (Est.=1.19, SE=0.07, t=15.50, p<.001) and arguments (Est.=0.70, SE=0.07, t=9.14, p<.001).

EXTRACTED CATEGORY did not have a significant effect (Est.=0.03, SE=0.07, t=0.43, p>.6), but there was a highly significant interaction between GAP POSITION and EXTRACTED CATEGORY (Est.=0.48, SE=0.08, t=5.61, p<.001). In the rationale clause condition, adjunct extraction scored much lower than argument extraction (Est.=−0.51, SE=0.07, t=−7.25, p<.001). In the matrix clause condition, there was no significant difference (Est.=−0.03, SE=0.07, t=−0.43, p>.6).

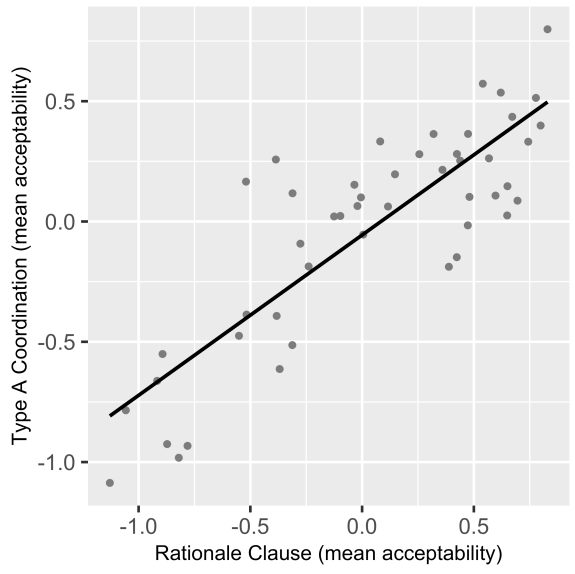


Figure 10: Correlation between the items in exp IVB and the items in exp. IVA

The acceptability of extraction from Type A coordination constructions and from rationale clause constructions showed a highly positive correlation (tau=0.5921986). About 70% of the variance of the acceptability of extraction from type A coordination can be explained based on the acceptability of extraction from rationale clauses (r^2=0.7078). This result is highly significant (t=10.556, df=46, p<0.001).

The observed highly significant correlation confirms prediction 8. As per predictions 6 and 6', adjunct extraction from the right domain in the structures under consideration scored much lower than argument extraction. As per predictions 7 and 7', adjunct and argument extraction from the left domain did not differ in acceptability. Finally, as predicted, argument extraction from the left domain was more acceptable than argument extraction from the right domain.

Thus, like rationale clauses, the right conjunct of a type A coordination is a selective island. By contrast, extraction from the left conjunct is like extraction from a matrix VP in not showing any island effects.

4.5.2 Experiment V

Our interpretation of the results of experiments IVA and IVB is based on the notion that VP-internal adjuncts are selective islands. There is a potential alternative explanation. In acceptability judgment experiments, length of movement is inversely correlated with acceptability (Sprouse et al. 2013, 2016; Kush et al. 2018). It is conceivable that this effect is stronger for adjuncts than for arguments. If so, the results of experiments IVA and IVB could be attributed to this interaction.

We therefore ran a final experiment to test whether the length effect is indeed stronger for adjuncts than for arguments, using clause-internal vs. cross-clausal movement. We expected that the length effect does not discriminate between arguments and adjuncts:

- Both argument and adjunct extraction are more degraded when they cross a clause boundary than when they do not (PREDICTION 9).
- Nonisland clausal boundaries reduce the acceptability of adjunct and argument extraction to the same degree (PREDICTION 10).

If these predictions are correct, the island effect in experiments IVA and IVB must be real.

The experiment had a 2×2 factorial design in which we manipulated LENGTH (clause-internal movement vs. cross-clausal movement) and EXTRACTED CATEGORY (argument vs. adjunct). We created twelve sets of four test items, distributed across four lists in Latin Square fashion. Each list also contained four practice items and twenty-four fillers. Nonpractice items were pseudorandomized per participant. Sample test items are given in 41.

(41) [Ali defeated the heavyweight champion with a well-timed right hook.]
  a. Kim said that it was the heavyweight champion that Ali defeated with a well-timed right hook. (Short | Argument)

b. It was the heavyweight champion that Kim said that Ali defeated with a well-timed right hook. (Long | Argument)
c. Kim said that it was with a well-timed right hook that Ali defeated the heavyweight champion. (Short | Adjunct)
d. It was with a well-timed right hook that Kim said that Ali defeated the heavyweight champion. (Long | Adjunct)

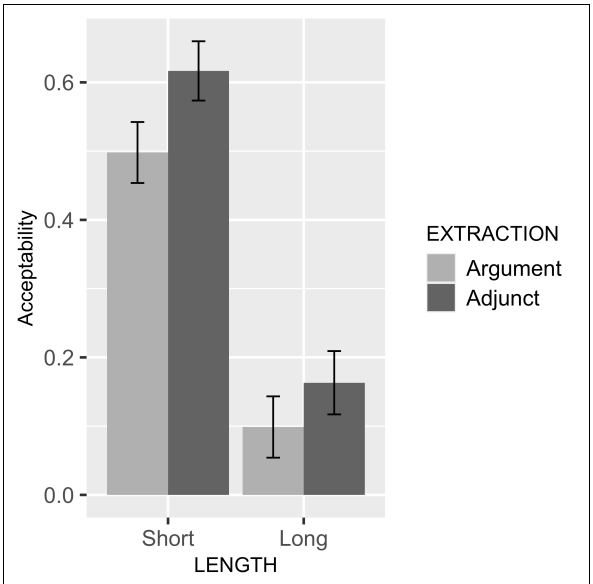


Figure 11: Mean acceptability by condition (z-scores; LENGTH × EXTR. CATEGORY) in exp. V

A linear mixed effects model was fitted to the results of the experiment, with LENGTH, EXTRACTED CATEGORY, and their interaction as fixed effect predictors. It showed a significant main effect of LENGTH (Est.=−0.39, SE=0.07, t=−5.58, p <.001): long movement came at a cost for both argument and adjunct extraction (arg: Est.=−0.39, SE=0.07, t=−5.58, p<.001; adj: Est.=−0.45, SE=0.07, t=−6.31, p<.001).

The effect of EXTRACTED CATEGORY was marginal (Est.=0.11, SE=0.06, t=1.96, p<.1). While adjunct clefting may be slightly easier than argument clefting, the contrast did not reach significance.

No interaction was found (Est.=−0.05, SE=0.08, t=−0.65, p>.5): long movement reduces the acceptability of argument and adjunct extraction equally. Indeed, in the long movement condition there was no difference between the two (Est.=−0.06, SE= 0.06, t=−1.06, p>.2).

Exp. V thus confirmed predictions 9 and 10: longer movements were rated lower, but this effect did not discriminate between argument and adjunct extraction. Hence, the results of experiments IVA and IVB cannot plausibly be ascribed to a discriminating length effect. They rather show that the right extraction domains in these experiments (right conjunct and rationale clause) are selective islands.

4.6 General discussion

We now return to the question whether CSC exceptions like type A coordination should receive a syntactic or an interpretive account. Our first claim:

A. There is an empirical parallel between extraction from structures containing rationale clauses and extraction from type A coordinate structures.

The interpretive accounts in Deane 1991, Na & Huck 1992, and Kehler 2002 and the syntactic accounts in De Vos 2005; 2009, Weisser 2015, and Bošković 2020 incorrectly predict that extraction from the *left* conjunct is harder than extraction from the *right* conjunct. Our second and third claims:

B. Like rationale clauses, the apparent right conjunct in a type A coordination is an adjunct.

C. Like transparent rationale clauses, the apparent right conjunct in a type A coordination is an adjunct attached VP-internally.

Claim C is not the only option. Altshuler and Truswell (2022) assume that categories adjoined to VP can be transparent for extraction of arguments. This allows them to analyze type A coordination as adjunction of the second conjunct *to* (rather than *within*) the first conjunct.

Should the acceptability of extraction from type A coordination be explained by assuming a structure distinct from regular coordination or by relying on its asymmetric interpretation?

- Claim C necessitates a syntactic account of type A coordination.
- Claim B is compatible with an interpretive account of type A coordination (if claim C is rejected). Munn (1993) proposes an analysis of regular coordination as rightward adjunction (see also Bošković & Franks 2000, Hartmann 2000, and Zhang 2010). Hence, if Munn is right, an interpretive explanation would be

required for CSC exceptions, as in Altshuler & Truswell 2022.

Acceptance of claim B under rejection of claim C is also compatible with a syntactic account of type A coordination. If type A coordination receives a Munn-style analysis, regular coordination must be assigned one of various competing analysis.

Thus, if claim C turns out to be incorrect, while claim B is confirmed, the choice is between a syntactic account of type A coordination and an interpretive account in which both type A coordination and regular coordination are adjunction structures. Which account is right must then be decided based on the properties of regular coordination. We have argued in section 2 that a Munn-style analysis of coordination cannot be correct.

- If both claim B and claim C are rejected, so that type A coordination is not adjunction, the empirical parallel with rationale clauses must presumably be explained in terms of the shared semantic properties of the two structures. Currently, no theories are available that can capitalize on this parallel, so a purely semantic account must remain a promissory note.

5. Compactness is linear

The OV/VO parameter seems to be connected the availability of scrambling, creating a pattern (word order is freer before the verb than after it) that contrasts with the Universal 20 generalization (word order is rigid before the noun and variable after it). We propose an analysis based on three core assumption three assumptions: (i) a separation of the argumental and adverbial hierarchies, (ii) a constraint on linearization in the VP, and (iii) a version of case adjacency.

This section is based on Janke, V., and A. Neeleman (2012). Ascending and Descending VPs in English. *Linguistic Inquiry* 43: 151-190; Neeleman, A., and A. Payne (2020a). PP Extraposition and the Order of Adverbials in English. *Linguistic Inquiry* 51: 471–520; and Neeleman, A., and A. Payne (2020b). On Matrix-Clause Intervention in Accusative-and-Infinitive Constructions. *Syntax* 23: 1–41.

5.1 The problem

The aim of this final section is to develop an analysis of VP in which certain differences between Dutch and English fall out from the XV/VX parameter. Starting point is the observation that Dutch allows adverbs to surface between the verb and its internal arguments, while such adverbial intervention is prohibited in English:

(1) a. Ava heeft ⟨snel⟩ Carlos ⟨snel⟩ het boek ⟨snel⟩ gegeven ⟨*snel⟩. *Dutch*
Ava has quickly Carlos quickly the book quickly given quickly
b. Ava has ⟨quickly⟩ given ⟨*quickly⟩ Carlos ⟨*quickly⟩ the book ⟨quickly⟩.

The contrast extends beyond Dutch and English. All known SOV languages with preverbal adverbs allow adverbial intervention. The picture for SVO languages is less clear, as verb movement may result in S-V-Adv-O order. However, on closer inspection *compactness* (to use Haider's (2005, 2014) term) is well motivated for SVO languages. Suppose that the verb moves and there are no adverbial positions between its trace and the object. Then, if two adverbs intervene between the verb and the object, the higher adverb must precede the lower adverb. By contrast, if there were adverbial positions between V and O, an intervening higher adverb should be able to *follow* an intervening lower one, whether the verb moves or not:

(2) a. [S [V [Adv_2 [Adv_1 [t_V O]]]]]
b. *[S [[[V Adv_1] Adv_2] O]]

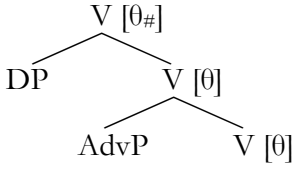
The following data thus confirm that Czech, even though it permits adverbial intervention, has a compact VP:

(3) a. Petr si přečetl ⟨znovu⟩ třikrát ⟨*znovu⟩ dopis. *Czech* (Jiri Kaspar, p.c.)
Peter.NOM REFL read again three-times again letter.ACC
'Peter again read the letter three times.' (*again* > *three times*)
b. Petr políbil ⟨?včera⟩ pomalu ⟨*včera⟩ Marii.
Peter.NOM kissed yesterday slowly yesterday Mary.ACC
'Yesterday Peter kissed Mary slowly.' (*yesterday* > *slowly*)

In my assessment, the contrast between Dutch and English extends to many other SOV and SVO languages and is at least correct as a statistical approximation.

The Dutch pattern is easy to explain under the proposal that syntactic dependencies are established through percolation and satisfaction of a selectional requirement. An analysis as in (4) captures core properties of neutral scrambling in Dutch, in particular the fact that it has A-properties (Huybregts and Van Riemsdijk 1985, Vanden Wyngaerd 1989, Neeleman 1994), does not reconstruct for scope (VandenWyngaerd 1989, Zwart 1993, Neeleman 1994) and does not lead to reactivation of the scrambled category in cross-modal priming experiments (Van de Koot et al. 2015).

(4)



A semantics of thematic selection compatible with this view of scrambling has recently been developed by Büring (2024).

Thus, in my view, the issue is how to account for the data in SVO languages. This is all the more urgent because the overall pattern seems to be the opposite of what we saw in Universal 20: there is word order freedom to the left of the head and strict work order following it.

5.2 Haven't we already solved this?

There are two traditional approaches to the English pattern. The first assumes that the verb and its internal arguments form a unit that excludes all other material, including adverbs (*impenetrability*; Chomsky 1965, 1986). The second approach assumes that the verb must be adjacent to DP objects (*case adjacency*; Keyser 1968, Stowell 1981). Both impenetrability and case adjacency rule out adverbial intervention, but…

- Neither impenetrability nor case adjacency sheds light on the grammaticality of adverbial intervention in Dutch (which is then dealt with through a movement operation absent in English, as in Kerstens 1975, De Haan 1979, Hoekstra 1984, …).
- Neither impenetrability nor case adjacency can cope with the distribution of very low adverbs, which must follow the object in English (Jackendoff 1972).
  - Impenetrability does not allow any adverb incursion into the VP (but such incursion exists).
  - Case adjacency only cares about adverbial intervention. It hence allows low attachment of adverbs whether they precede the verb or follow the object (but *[[Adv V] DP] is ungrammatical).

(5) a. Ava heeft ⟨*slecht⟩ Carlos ⟨slecht⟩ behandeld.
Ava has badly Carlos badly treated
'Ava has treated Carlos badly.'
b. Ava ⟨*badly⟩ treated Carlos ⟨badly⟩.

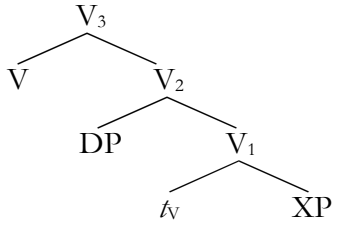
(6) a. Ava heeft haar kinderen elegant aangekleed. (manner; quality)
Ava has her children elegantly dressed
'Ava has dressed her children elegantly.'
b. Ava heeft elegant haar kinderen aangekleed. (manner; *quality)
Ava has eleganty her children dressed
'Ava has elegantly dressed her children.'
c. Ava dressed her children elegantly. (manner; quality)
d. Ava elegantly dressed her children. (manner; *quality)

(7) [Things are going downhill with John's company. On Monday, he had to lay off two employees. Unfortunately, that turned out not to be enough. (again > 2)].
a. Op dinsdag heeft Jan ⟨opnieuw⟩ twee medewerkers ⟨#opnieuw⟩ ontslagen.
on Tuesday has John again two employees again fired
b. On Tuesday John ⟨again⟩ fired two employees ⟨again⟩.

(8) [Things are going downhill with John's company. On Monday, he had to lay off ten employees. Unfortunately, he forgot to record who exactly he had fired. The next day, it became clear that another ten lay-offs were needed. Embarrassingly, John called up some people he had already fired. (2 > *again*)].
a. Op dinsdag heeft Jan ⟨#opnieuw⟩ twee medewerkers ⟨opnieuw⟩ ontslagen.
on Tuesday has John again two employees again fired
b. On Tuesday John ⟨#again⟩ fired two employees ⟨again⟩.

There is consensus that low adverbs should be analyzed as occupying the lowest position in a VP-shell structure (XP in (9)) (Larson 1988, Vanden Wyngaerd 1989). If the verb always moves, it also follows that preverbal adverbs are attached high. However, as verb movement in general can cross adverbs, it is unclear why no adverb can be left-adjoined to either V_1 or V_2, as shown in (10).

(9)



(10) a. *Ava dressed [$_{V2}$ her children [$_{V1}$ twice [$_{V1}$ t_V elegantly]]].
b. *Ava dressed [$_{V2}$ twice [$_{V2}$ her children [$_{V1}$ t_V elegantly]]].
c. *Ava has given [$_{V2}$ Carlos [$_{V1}$ quickly [$_{V1}$ t_V the book]]].
d. *Ava has given [$_{V2}$ quickly [$_{V2}$ Carlos [$_{V1}$ t_V the book]]].

- Verb movement plus case adjacency straightforwardly rules out (10b) and (10d) (Chomsky 1995, Janke and Neeleman 2012). However, case adjacency has nothing to say about (10a) (*elegantly* does not bear case).
- Haider (2005, 2014) proposed an adjusted version of impenetrability that bans attachment of elements in the structure between V and t_V (based on the idea that selection requires derivational sisterhood). However, floating quantifier can appear VP internally (Maling 1976, Baltin 1995, Janke and Neeleman 2012). If floating quantifiers are adjuncts (Kayne 1975, Dowty and Brody 1984), this poses a straight-forward problem.

(11) a. Ava dressed [$_{V2}$ her children [$_{V1}$ both [$_{V1}$ [$_{V1}$ t_V elegantly] on two occasions]].
b. Ava gave [$_{V2}$ the boys [$_{V1}$ each [$_{V1}$ [$_{V1}$ t_V a bag of sweets] twice]].

5.3 A proposal

The proposal is based on three assumptions: (i) separate argumental and adverbial hierarchies, (ii) a constraint on linearization in the VP, and (iii) a version of case adjacency.

5.3.1 Adverbial intervention

Consider a Dutch example in which a triadic verb appears in a structure with a temporal and a manner adverb. Of the 120 possible orderings of the five preverbal constituents, only ten are grammatical (in the absence of contrast):

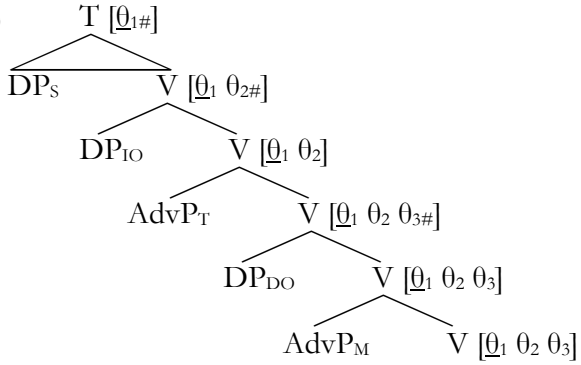
(12) a. Volgens mij heeft toen ⟨snel⟩ Ava ⟨snel⟩ de jongens ⟨snel⟩ het boek ⟨snel⟩ gegeven.
according.to me has then quickly Ava quickly the boys quickly the book quickly given
'I think that Ava quickly gave the boys the book at that point.'
b. Volgens mij heeft Ava toen ⟨snel⟩ de jongens ⟨snel⟩ het boek ⟨snel⟩ gegeven.
according.to me has Ava then quickly the boys quickly the book quickly given
c. Volgens mij heeft Ava de jongens toen ⟨snel⟩ het boek ⟨snel⟩ gegeven.
according.to me has Ava the boys then quickly the book quickly given
d. Volgens mij heeft Ava de jongens het boek toen snel gegeven.
according.to me has Ava the boys the book then quickly given

Given that the account of Universal 20 yields a strict pre-head order in the case of a single hierarchy of elements, the pattern in (12) requires that there are separate hierarchies that respectively regulate the projection of arguments and the attachment of adverbs (as argued by Bobaljik 1999).

(13) a. S > IO > DO
b. Time adverb > Manner adverb

The adverbial hierarchy is presumably grounded in semantics (Ernst 2002, Nilsen 2003). The argumental hierarchy presumably results from a (semantically motivated) ordering of θ-roles in the verb (Levin and Rappaport Hovav 2005). Thus, (12c), with 'quickly' in the lower position, can be assigned the following structure:

(14)



5.3.2 Canonical Linearization

I propose that the parameter that distinguishes SOV and SVO languages determines linearization in the verb's selection domain:

(15) *XV/VX parameter*
   a. Within V's selection domain, a node V_i is ordered canonically if the category from which it is projected {precedes, follows} its sister.
   b. V's selection domain is the minimal subtree that contains all categories selected or licensed by V under projection.

The effects of (15) come about through various ordering statements that mention or are sensitive to canonical order. The simplest such statement directly mentions the categories to be ordered, as in (16).

(16) Any subtree $[_{V2}\ V_1\ XP]$ where XP ∈ {AdvP, PP, AP, …} is ordered canonically.

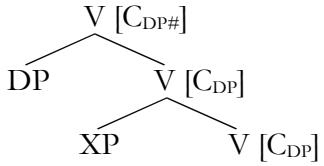
As predicted, there is a strong typological correlation between the order of verb and object, the placement of (low) adverbs and the placement of adpositional phrases (data in (17) are from Dryer 1991):

(17)

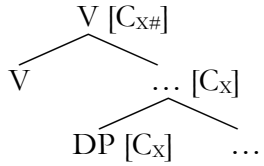
	SOV	SVO
Manner adverb – V	0.91	0.25
Adpositional phrase – V	0.90	0.01

DP arguments are not ordered by (16). They are special in having separate systems for selection and licensing. Selection is mediated by θ-theory, while licensing is the subject of case theory. Case licensing is always upward but can be initiated by either the licensing head or the case-marked DP (cf. Ackema and Neeleman 2018). In the first case, the DP must c-command the head, in the second case, the head must c-command the DP.

(18) a.



b.



The linearization rule for DPs refers to case: nodes in which C_{DP} or C_X are satisfied must be ordered canonically. This is a reformulation of the traditional claim that SVO and SOV languages differ in the direction of case assignment.

(19) Any subtree $[_{V2}\ V_1\ XP]$ in which V_2 contains either $C_{DP\#}$ or $C_{X\#}$ is ordered canonically.

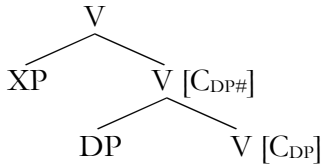
I will refer to the combination of the XV/VX parameter and the constraints sensitive to it as *Canonical Linearization* (CL).

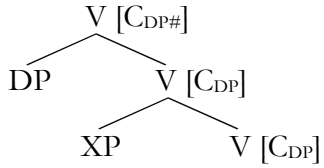
5.3.3 The Case-First Constraint

This brings us to the second constraint central to the analysis of compactness – a reformulation of Stowell's (1981) notion of Case Adjacency. However, rather than insisting on adjacency per se, (20) requires that no phrase precede a case-marked argument in the string that runs from that argument to the case-marking head.

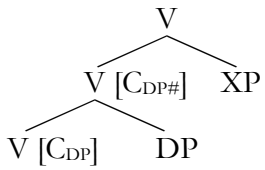
(20) *Case-First Constraint* (CFC)
   a. A DP argument must be the leftmost phrase in its case domain.
   b. The case domain of a DP argument consists of that DP and any material that linearly intervenes between it and its case licenser.

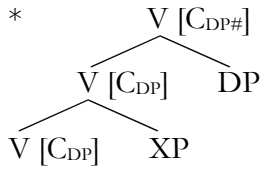
This has the advantage that the CFC can apply to SOV and SVO languages alike. In SOV languages it allows the flexibility inherent in the system to surface. In (21a), the DP's case domain is ⟨DP⟩, so the DP comes first. The same is true of the DP's case domain in (21b), which is ⟨DP, XP⟩. Thus, both structures are ruled in.

(21) a.


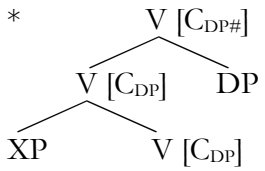
b.


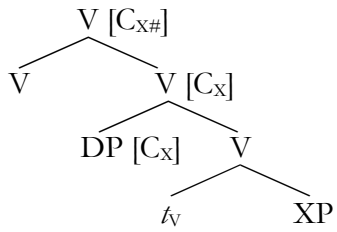
In SVO languages, the CFC obscures the flexibility inherent in the system. (21a) has a direct counterpart in (22a). (21b), however, has no direct counterpart. (22b) violates the CFC (the DP's case domain is ⟨XP DP⟩).

(22) a.


b. *


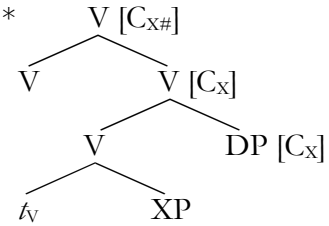
The linearization problem in (22b) cannot be fixed by simply pivoting the XP out of the way. (23a) does not violate the CFC, but it is still ruled out as the linearization of XP is not harmonic (see also section 4).

(23) a. *


b.


VP-shell structure in (23b), however, satisfied both CL and the CFC.

- Linearization is canonical, even though the DP in (23b) precedes the verb's base position. DP objects are subject to (19) (rather than to (16)). Hence, the top node in (23b) must be linearized canonically. However, no instance of C is satisfied in the node that immediately dominates the DP, and so the DP can be linearized to the left of t_V.
- The structure satisfies the CFC. As the DP's case is licensed by the verb in its derived position, the DP's case domain consists of just the DP, and hence it is leftmost in its case domain, as required.
- Note that CL allows the DP to be placed to the right of t_V. However, that linearization violates the CFC. The DP's case domain is the string ⟨XP, DP⟩, in which the DP is preceded by XP.

(24) *


Two further analytical details need to be added:

- Verb movement is self-attachment (Ackema, Neeleman, and Weerman 1993, Koeneman 2000, Hornstein and Uriagereka 2002, Bury 2003, Fanselow 2003, Surányi 2005 and Bayer and Brandner 2007). This explains why verb movement in (23b) is obligatory. C_X needs to be satisfied by a case licensing head, which will be absent if the verb does not move.

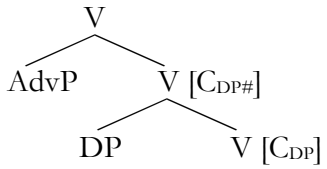
- Case licensing through C_X is dispreferred as compared to case licensing through C_{DP}. Hence, a DP that introduces C_X can only be merged in positions where case licensing through C_{DP} is blocked. No VP-shell formation will take place unless a category is attached to the right of V before the DP is attached.

5.4 The way the English VP bends

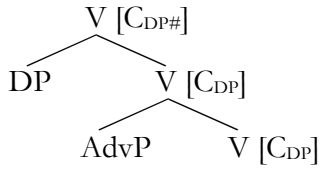
5.4.1 DP complements

The system outlined above captures the data discussed in sections 1 and 2. It implies that in Dutch the attachment height of adverbs is reflected by (preverbal) linear order:

(25) a.

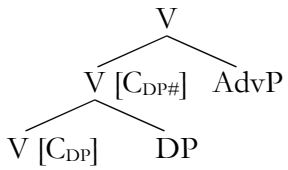


b.

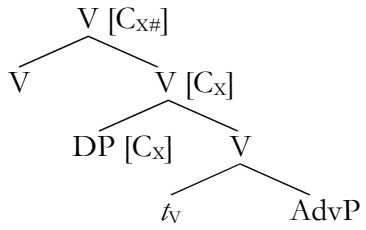


In English, structural intervention is possible, but the CFC forces VP-shell formation when a DP argument is attached higher than an adverb, so that structural intervention does not lead to linear intervention (Larson 1988, Vanden Wyngaerd 1989):

(26) a.



b.



Evidence for (26a) comes from the stranding of adverbs under VP-fronting and VP-ellipsis (among other things):

(27) a. Ava wanted to read the book, and [$_V$ read the book] she did yesterday.
b. Ava wanted to read the book, and she did so yesterday.

The same constituency tests confirm (26b). Adverbs whose counterpart in Dutch must be attached lower than the object, cannot be stranded under VP-fronting and VP-ellipsis:

(28) Ava treated [Carlos [t_V badly]].

(29) a. *Ava wanted to treat Carlos like Carlos had treated her, so [treat Carlos] she did badly.
b. *Ava wanted to treat Carlos like Carlos had treated her, so she did so badly.

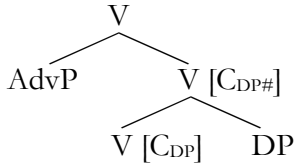
(30) a. Ava [[dressed her children] elegantly]. (manner)
b. Ava [dressed [her children [t_V elegantly]]. (quality)

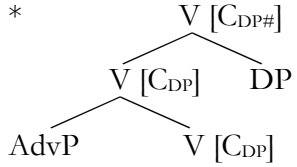
(31) a. Ava wanted to dress her children like she dresses herself,
so [$_V$ dress her children] she did elegantly. (manner; *quality)
b. Ava wanted to dress her children like she dresses herself,
so she did so elegantly. (manner; *quality)

(32) a. John [[fired two employees] again]. (*again* > *2*)
b. John [fired [two employees [t_V again]]. (*2* > *again*)

(33) a. He fired two employees on Monday,
and [$_V$ fire two employees] he did again on Tuesday. (*again* > *2*; **2* > *again*)
b. He fired two employees on Monday,
and on Tuesday he did so again. (*again* > *2*; **2* > *again*)

A second consequence of the proposal is that preverbal adverbs must be attached outside the verb's selection domain, as within the selection domain CL requires postverbal linearization:

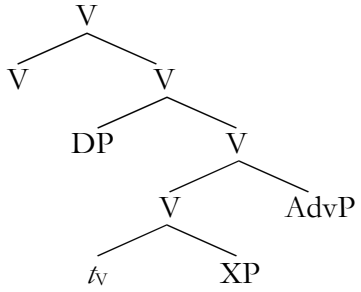
(34) a. 

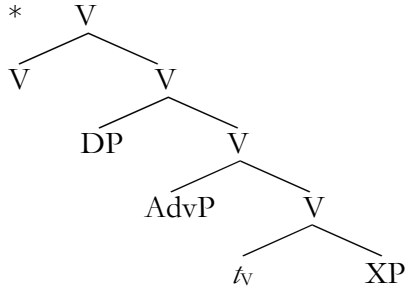
b. * 

This captures the generalization that there is incursion into the English VP by right-attached, but not by left-attached adverbs (see the discussion around (5b), (6c)/(6d) or (7b)/(8b)):

(35) a. *Ava badly [treated Carlos].
b. Ava elegantly [dressed her children]. (manner; *quality)
c. John again [fired two employees]. (*again* > *2*; **2* > *again*)

CL has a further consequence. Suppose that an adverb is attached within the lower part of a VP-shell, but higher than the lowest position in the structure. As such an adverb is part of the verb's selection domain, it must be linearized to the right of its sister. This is the third key observation from section 2, where I also gave some evidence from floating quantifiers that the structure in (36a) exists.

(36) a. 

b. * 

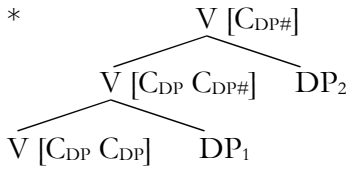
(37) a. Ava heeft Carlos ⟨twee keer⟩ <u>slecht</u> ⟨*twee keer⟩ behandeld.
Ava has Carlos two times badly two times treated
'Ava has treated Carlos badly twice.'
b. Ava heeft haar kinderen ⟨twee keer⟩ <u>elegant</u> ⟨*twee keer⟩ aangekleed. (quality)
Ava has her children two times elegantly two times dressed
c. Jan heeft twee werknemers ⟨zojuist⟩ <u>opnieuw</u> ⟨*zojuist⟩ ontslagen. (*2* > *again*)
John has two employees just.now again just.now fired
'John has fired two employees again just now.'

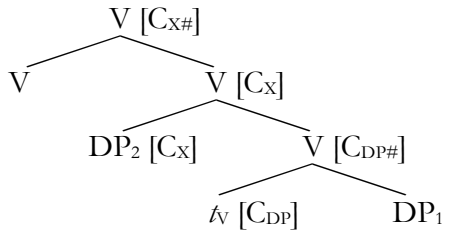
(38) a. Ava treated Carlos ⟨*twice⟩ <u>badly</u> ⟨twice⟩.
b. Ava dressed her children ⟨*twice⟩ <u>elegantly</u> ⟨twice⟩. (quality)
c. John fired two employees ⟨*just now⟩ <u>again</u> ⟨just now⟩. (*2* > *again*)

The analysis extends straightforwardly to double-object and dative constructions. The Dutch double-object construction has strict indirect object–direct object order. This suggests that in the verb's θ-grid the θ-role assigned to the direct object is subordinate to the one assigned to the indirect object, as in (14) (this should be true in English as well).

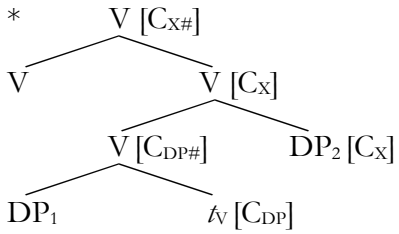
(39) Ava heeft ⟨Carlos⟩ het boek ⟨??Carlos⟩ gegeven.
Ava has Carlos the book Carlos given
'Ava has given Carlos the book.'

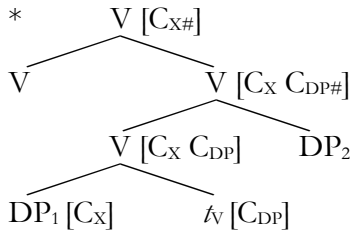
In English, VP-shell formation is necessary). (40a) violates the CFC, but (40b) is grammatical. The CFC is satisfied if V case-marks DP_2 while t_V case-marks DP_1. CL is satisfied, as the nodes specified [$C_{X\#}$] and [C_{DP}#] are linearized canonically.

(40) a. * 

b. 

A structure with DP_1 adjacent to the verb and DP_2 adjacent to the verb's trace violates both CL and the CFC if the objects are licensed as in (41b). Hence, DP_1 must be licensed by the verb in its derived position, while DP_2 is licensed by the verb's trace. Recall, however, that a DP that introduces C_X can only be merged in positions where case licensing through C_{DP} is blocked. As DP_1 can in fact be licensed through C_{DP}, (41b) is ruled out as well.

(41) a. * 

b. * 

The upshot is that English double-object constructions, too, have a strict indirect object–direct object order:

(42) Ava gave ⟨Carlos⟩ the book ⟨*Carlos⟩

The structure in (40b) is of course not controversial. To begin with, in examples like (43) the direct object appears in the scope of the indirect object (Larson 1990, Bruening 2001, Heizmann 2007, Bruening 2019).

(43) I gave a student every book. (∃ > ∀; *∀ > ∃)

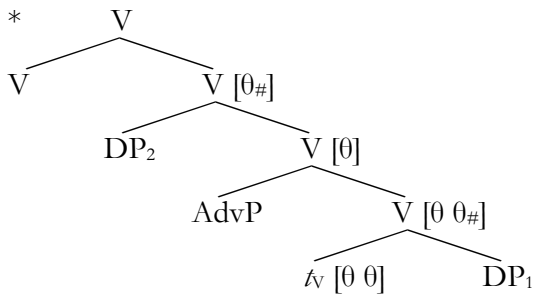
Furthermore, the V-DP_2 string does not meet constituency tests:

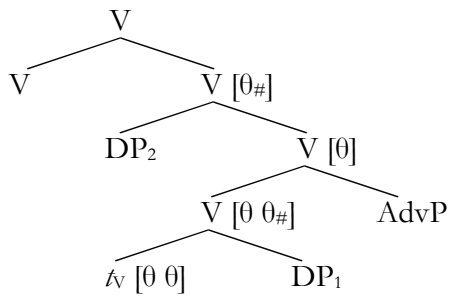
(44) a. *Ava wanted to [give Carlos something comforting], so give Carlos she did a woolen scarf.
b. *If Ava [gave Carlos anything], she did so a woolen scarf.
c. [Give Carlos ⟨the books⟩] though she may ⟨*the books⟩, it won't make a difference.

It is predicted that adverbs can appear within the lower part of the shell structure, but they cannot linearly intervene between the verb and either of its objects. This yields the pattern that we started out with in (1):

(45) Ava has ⟨quickly⟩$_1$ given ⟨*quickly⟩$_2$ Carlos ⟨*quickly⟩$_3$ the book ⟨quickly⟩$_4$.

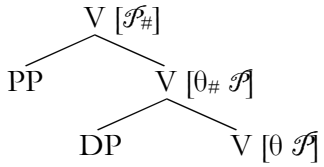
Position 1 is available if the adverb is outside the verb's selection domain. Position 2 violates the CFC. Position 3 violates (16) (see (46a)). Position 4 is grammatical and should be parsable with the adverb attached between the two objects (see (46b)).

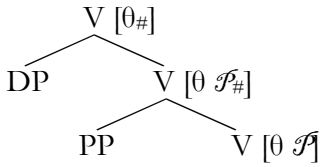
(46) a. * 

b. 

In Dutch dative constructions, word order is variable. It is not clear to me why this should be, but it suggests that the selectional requirement that licenses PP complements is not (necessarily) ordered with respect to θ-roles:

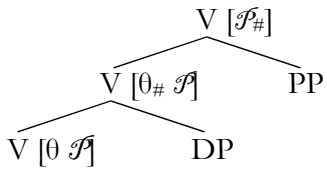
(47) Ava heeft ⟨aan Carlos⟩ de krant ⟨aan Carlos⟩ gegeven.
Ava has to Carlos the newspaper to Carlos given
'Ava has given the newspaper to Carlos.'

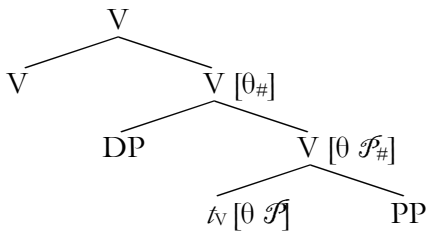
(48) a.


b.


In English dative constructions, word order is strict, but an analysis on a par with Dutch is possible given that the CFC triggers VP-shell formation:

(49) Ava gave ⟨*to Carlos⟩ the newspaper ⟨to Carlos⟩.

(50) a.


b.


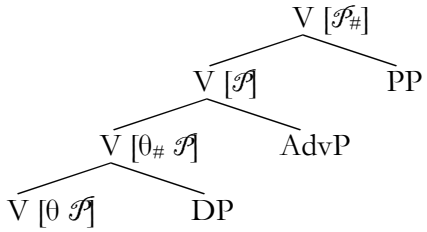
Given the structural ambiguity in (50), scope should be variable, and the V-DP string should meet constituency tests:

(51) a. I read a book to every child. (∃ > ∀ <> ∀ > ∃)
b. I read every book to a child. (∃ > ∀ <> ∀ > ∃)

(52) a. Ava wanted to read the poems to someone she liked, and [read the poems] she did to Carlos.
b. If Ava [read the poems] to anyone, she did so to Carlos.
c. [Give the books] though she may to Carlos, it won't make a difference.

It is also predicted that placement of adverbs in the dative construction is freer than in the double object construction. In particular, position 3 should be available on an ascending analysis of the VP (compare (45)):

(53) Ava ⟨quickly⟩$_1$ gave ⟨*quickly⟩$_2$ the newspaper ⟨quickly⟩$_3$ to Carlos ⟨quickly⟩$_4$.

(54)


5.4.2 Floating quantifiers

The distribution of floating quantifiers (FQs) supports aspects of the analysis developed above. In both Dutch and English, FQs can be analyzed as adjuncts that are linked to an unassigned θ-role (cf. Baltin 1978, 1982, 1995, Belletti 1982, Bobaljik 1995, Doetjes 1997 and Janke and Neeleman 2012). They are not subject to (16) or (19) but are linearized *between* the recipient of that θ-role and the predicate that introduces it.

(55) a. FQs associate with an unassigned θ-role in the node they combine with.
b. FQ^i is linearized between the predicate that introduces θ^i and the DP that satisfies θ^i.

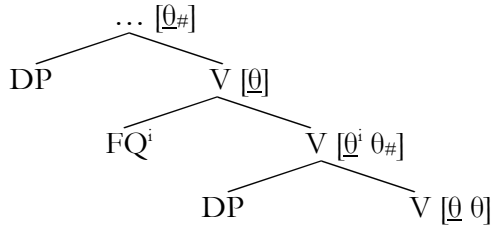
The effects of these requirements in Dutch are straightforwardly. In both (56a) and (56b), position 1 violates both (55a) and (55b), position 2 meets both (55a) and (55b), and position 3 meets (55a) but not (55b).

(56) a. Op maandag hebben ⟨*elk⟩$_1$ de jongens ⟨elk⟩$_2$ hun examen gedaan ⟨*elk⟩$_3$.
on Monday have each the boys each their exam done each
'On Monday, each of the boys sat their exam.'

b. Op maandag heb ik ⟨*elk⟩$_1$ de jongens ⟨elk⟩$_2$ gefeliciteerd ⟨*elk⟩$_3$.
on Monday have I each the boys each congratulated each
'On Monday, I congratulated each of the boys.'

The example in (56a) is analyzed as below (I am ignoring possible A-movement to the subject position; see Williams 1986 and Neeleman and Van de Koot 2010 for an analysis of A-trace as introducing a θ-role satisfied by the raised category).

(57)



In English, the constraints in (55) are unproblematic for subject-oriented FQs: position 1 in (58) violates both clauses, position 2 satisfies both, and position 3 satisfies (55a), but not (55b).

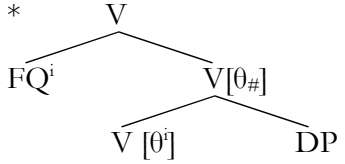
(58) On Monday, ⟨*each⟩$_1$ the boys ⟨each⟩$_2$ sat the exam ⟨*each⟩$_3$.

What is striking is that English does not permit the insertion of object-oriented FQs in examples like (59) (see Maling 1976).

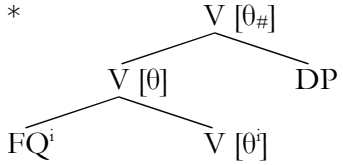
(59) *On Monday, I ⟨each⟩$_1$ congratulated the boys ⟨each⟩$_2$.

The ungrammaticality of (59) can be accounted for as follows. Position 1 violates both clauses in (55) if analyzed as in (65a), and (55b) if analyzed as in (65b).

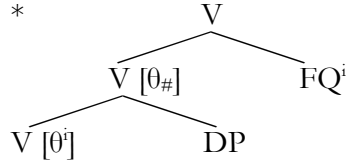
(60) a. *



b. *

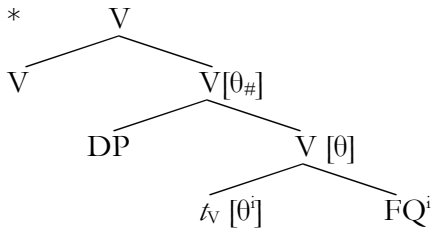


c. *

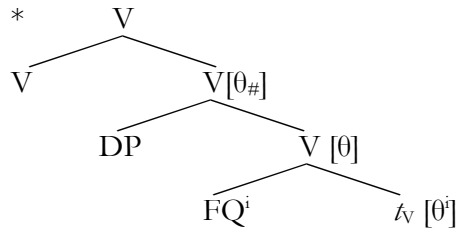


Position 2 violates both clauses in (55) if analyzed as in (60c) above. VP-shell formation does not help. In (61a), the FQ is linearized contrary to (55b). In (61b), there is no postverbal category and hence no license for shell formation (see section 3.3.3).

(61) a. *

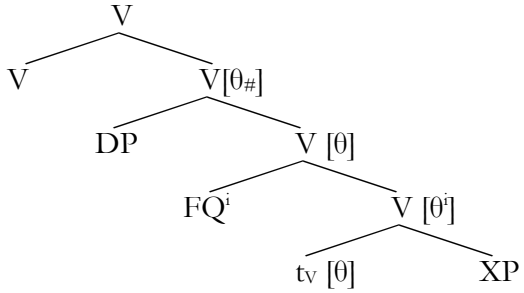


b. *

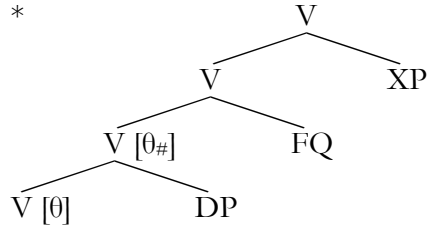


This analysis of the ungrammaticality of (59) makes a crucial prediction. An object-oriented FQ can be saved if some other element is attached before it in a VP-shell structure. In (62a), low right-attachment of the XP triggers VP-shell formation, while left attachment of the FQ satisfies (55b). Note that the saving effect of XP relies on it being attached low, in a VP-shell structure. If XP is part of an ascending structure, as in (62b), the FQ violates both clauses of (55).

(62) a.



b. *



Maling already observed that object-oriented FQs can be saved by certain elements that follow them. I first consider structures in which XP is an adverb. The saving effect is shown in (63):

(63) a. Ava read these papers all *(very carefully).
b. Carlos sliced the onions both *(quite thinly).

As predicted, adverbs that must be attached high cannot rescue object-oriented FQs:

(64) a. Ava ⟨probably⟩ didn't ⟨*probably⟩ read all these papers
b. *Ava didn't read these papers all, probably.
c. Carlos ⟨unfortunately⟩ didn't ⟨*unfortunately⟩ slice both the onions.
d. *Carlos didn't slice the onions both, unfortunately.

The effect is also observable when manner and time adverbs take on the role of savior (see Neeleman and Payne 2020a):

(65) a. Ava studied the letters both carefully.
b. ??Ava studied the letters both yesterday.

Constituency tests confirm that elements that can normally be stranded under VP-fronting or VP-ellipsis can no longer do so if they act as the savior of an object-oriented FQ.

(66) a. Ava promised she would read the two books I sent,
and read both books she did carefully.
b. Ava promised she would carefully read the two books I sent,
and read both books carefully she did.

(67) a. *Ava promised she would read the two books I sent,
and read the books <both> she did <both> carefully.
b. Ava promised she would carefully read the two books I sent,
and read the books both carefully she did.

Further evidence that FQs force a descending structure where an ascending one is normally available comes from temporal *only*. This element must immediately c-command its semantic argument (Barbiers 1995, Neeleman and Van de Koot 2021):

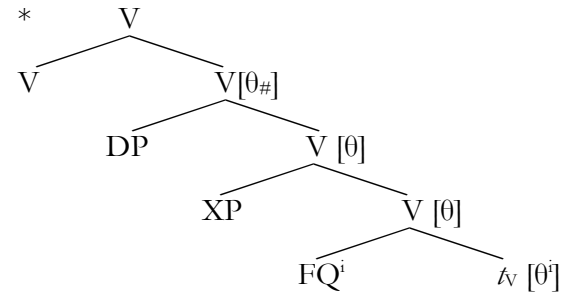
(68) Ava heeft pas ⟨*deze problemen⟩ op zondag ⟨deze problemen⟩ (allebei) opgelost.
Ava has only these problems on Sunday these problems both solved

This forces high attachment of the temporal PP in examples like (69b). At the same time, the FQ in (69c) forces low attachment. It is predicted, then, that (69c) should be ungrammatical.

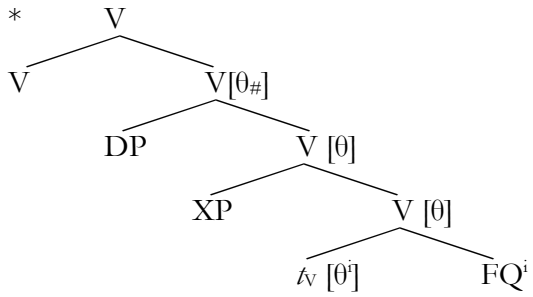
(69) a. Ava solved these problems both on Sunday.
b. Ava only solved these problems on Sunday.
c. *Ava only solved these problems both on Sunday.

The analysis predicts the adverbial savior must follow the floating quantifier. The structures in (70) are out for the same reason as those in (61), in addition to violating CL.

(70) a. *



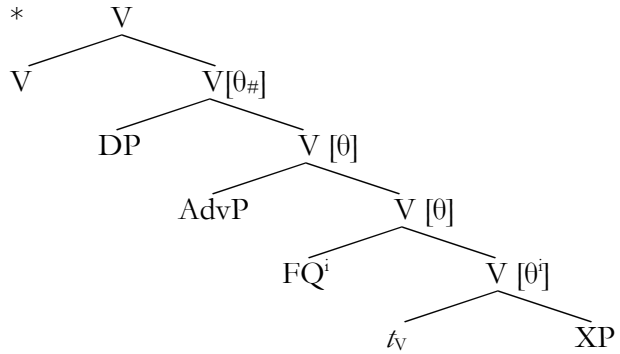
b. *



(71) a. *Moira read these papers very carefully all.
b. *I sliced the onions quite thinly both.

Similarly, even in the presence of a savior no adverb can surface between the object and the FQ. The node dominating AdvP in (72) violates CL.

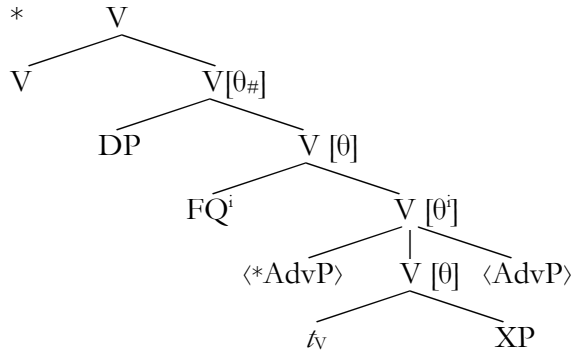
(72) *



(73) a. Moira read these papers ⟨*twice⟩ all very carefully ⟨twice⟩.
b. I sliced the onions ⟨*already⟩ both quite thinly ⟨already⟩.

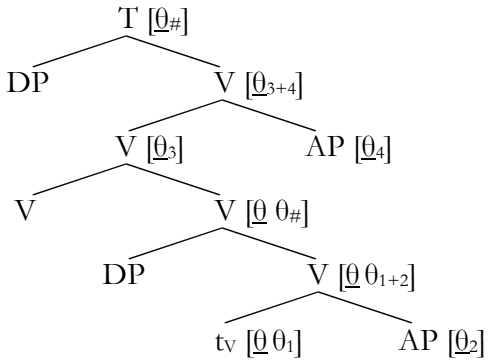
Although this is not an additional data point (cf. (36)), it is worth pointing out that adverbs following an object-oriented FQ must come in ascending order. The reason is familiar: left-attachment of AdvP in (74) violates CL.

(74) *



The argument can be replicated using secondary predicates. These are interesting to consider, as object-oriented predicates are attached below the object, while subject-oriented predicates (at least in English) are adjoined to VP:

(75)



(76) a. *The boys ate the meat drunk raw.
b. ?The boys ate the meat raw drunk.
c. *The boys painted the barn drunk green.
d. ?The boys painted the barn green drunk.

Hence, object-oriented, but not subject-oriented predicates should be able to rescue object-oriented FQs (see (62a,b), with XP a predicate). This turns out to be correct (Maling 1976):

(77) a. Ava photographed the boys both dressed in red.
b. Ava painted the doors both bright green.

(78) a. *Ava photographed the boys both dressed in red.
b. *Ava sat the exams both rather drunk.

Like adverbial saviors, predicates that save an object-oriented floating quantifier cannot precede that element (see (70a,b), with XP a predicate), while the FQ cannot be preceded by an adverb (see (62), with XP a predicate).

(79) a. *Ava photographed the boys dressed in red both.
b. *Ava painted the doors bright green both.

(80) a. Ava photographed the boys ⟨*twice⟩ (both) dressed in red ⟨twice⟩.
b. Ava painted the doors ⟨*already⟩ (both) bright green ⟨already⟩.

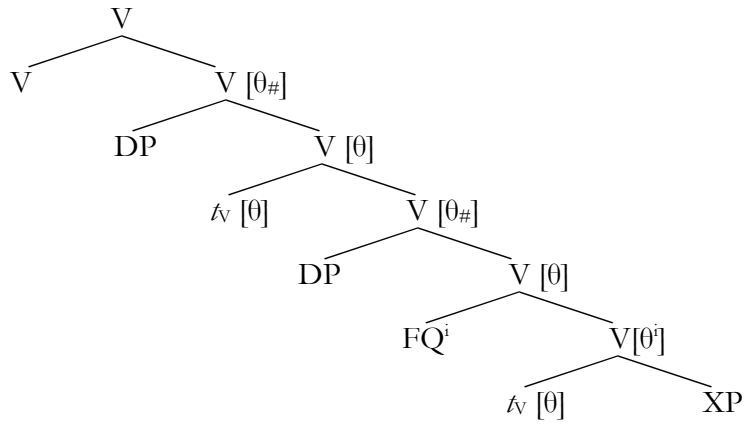
In Dutch double-object constructions, floating quantifiers can be associated with either object without difficulty:

(81) a. Ava heeft de jongens allemaal een boek gegeven.
Ava has the boys all a book given
'Ava had given all the boys a book'
b. Ava heeft Carlos de boeken allemaal gegeven.
Ava has Carlos the books all given
'Ava has given Carlos all the books.'

The prediction for English is that FQs can be associated with indirect objects as a matter of course (the direct object acts as savior; see (72) with XP a direct object). However, association with a direct object requires a double shell structure with an additional savior (as in (83)).

(82) a. Ava gave the boys both a good talking to.
b. *Ava gave Carlos the books both.
b. Ava gave Carlos the books both very quietly.
c. Ava gave Carlos the books both covered in dust.

(83)



In Dutch dative constructions, the DP object can be associated with an FQ irrespective of the position of the PP:

(84) a. Ava heeft de gedichten ⟨allebei⟩ aan Carlos ⟨allebei⟩ voorgelezen.
Ava has the poems both to Carlos both read
b. Ava heeft aan Carlos de gedichten allebei voorgelezen.
Ava has to Carlos the poems both read
'Ava read the poems both to Carlos.'

The prediction for English is that the PP argument can act as a savior for an object-oriented FQ, but only if merged low (see (62), with XP a dative PP). The effect was already observed in Maling 1976:

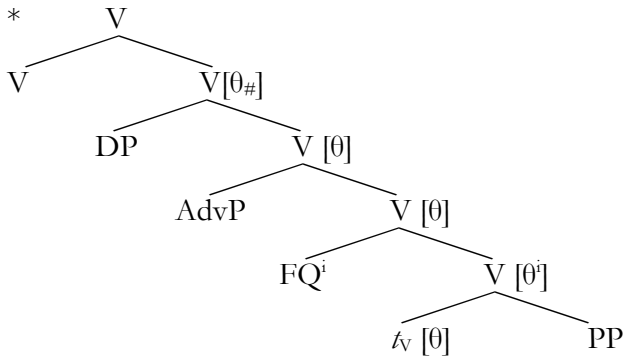
(85) Ava read the poems both *(to Carlos).

Since an FQ associated with the DP argument forces the PP in a low position, constituency tests yield different results depending on the presence or absence of an FQ:

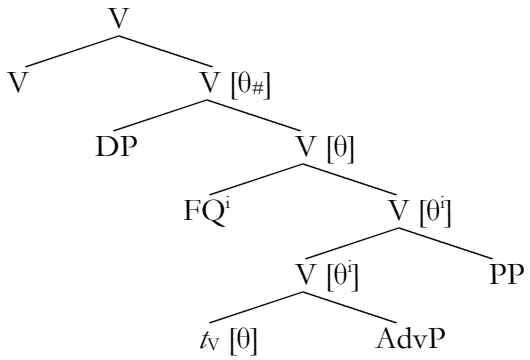
(86) a. If Ava read the poems to anyone, she did so (*both) to Carlos.
b. Ava wanted to read the poems to someone she liked,
and read the poems she did (*both) to Carlos.
c. Give the books (*both) though she may to Carlos, it won't make a difference.
d. Give the books though she may (*both) to Carlos, it won't make a difference.

As in shell structures, no adverb may appear between the DP argument and the FQ (because that would violate CL). However, an adverb *can* appear between the FQ and the PP. This is because the adverb can act as savior, while the PP is attached higher.

(87) a. *



b.



(88) Ava read the poems ⟨*quietly⟩ both ⟨quietly⟩ to Carlos.

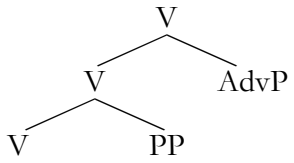
5.4.3 PP Complements

Like DP objects, PP complements in Dutch are freely ordered with respect to adverbs:

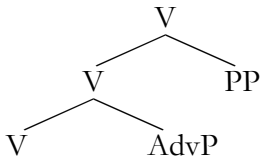
(89) Ava heeft ⟨twee weken lang⟩ aan dit probleem ⟨twee weken lang⟩ gewerkt.
Ava has two weeks long on this problem two weeks long worked
'Ava has worked on this problem for two weeks'

Given that the CFC does not apply to PPs, the alternation between Adv-PP-V and PP-Adv-V in (89) should have a direct counterpart in English (see also Stowell 1981):

(90) a.

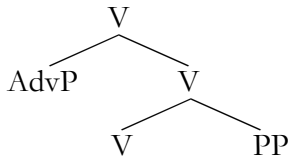


b.

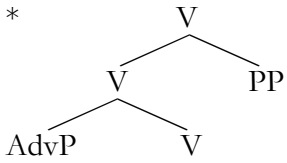


Of course, an adverb attached higher than the PP complement can also be attached to the left of the verbal projection line. Low adverbs, however, are subject to CL and must hence follow V.

(91) a.

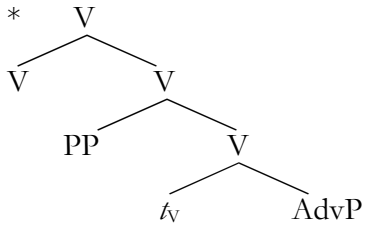


b. *



Finally, as VP-shell formation is case-driven operation, no shell formation is expected in structures with a PP complement.

(92) *



Given the ungrammaticality of (92), the following two predictions should hold:

i. Adverbs that follow a PP complement and preverbal adverbs c-command the PP complement, as in (90a) and (91a).
ii. A PP complement c-command adverbs sandwiched between the verb and the PP complement, as in (90b).

These predictions are borne out by the data in (93) and (94).

(93) [Participants in the workshop that Suzanne is organizing are unhappy about their accommodation. Two approached her with complaints this morning. She spoke with them and offered a partial refund. That settled things for them. However, in the afternoon two further participants approached Suzanne with complaints. (again > 2)]

a. 's Middags heeft Suzanne ⟨opnieuw⟩ met twee deelnemers ⟨#opnieuw⟩ gesproken.
in.the afternoon has Suzanne again with two participants again spoken

b. In the afternoon Suzanne ⟨again⟩ spoke ⟨#again⟩ with two participants ⟨again⟩.

(94) [Participants in the workshop that Suzanne is organizing are unhappy about the accommodation they were given. Ten approached her with complaints this morning. She spoke with them and offered a partial refund. That settled things for eight of the ten complainants, However, two were still unhappy. In the afternoon, a second group of ten participants approached Suzanne, which included the two still-unhappy participants that she spoke to in the morning. (2 > again)]

a. 's Middags heeft Suzanne ⟨#opnieuw⟩ met twee deelnemers ⟨opnieuw⟩ gesproken.
in.the afternoon has Suzanne again with two participants again spoken

b. In the afternoon Suzanne ⟨#again⟩ spoke ⟨again⟩ with two participants ⟨#again⟩.

The proposal also makes predictions for c-command between adverbs. For non-adjacent adverbs, these are given below (prediction (v) comes about because outside the verb's selection domain linearization of adverbs is variable, so that [$AdvP_1$ [VP $AdvP_2$]] and [[$AdvP_2$ VP] $AdvP_1$] are both permitted structures.)

iii. Preverbal adverbs c-command adverbs sandwiched between the verb and a PP complement.
iv. Adverbs that follow a PP complement c-command adverbs sandwiched between the verb and a PP complement.
v. Preverbal adverbs can either c-command or be c-commanded by adverbs that follow a PP complement.

These predictions appear to be correct, as already shown (for the most part) in Neeleman and Payne 2020. To begin with, consider *again* and *continuously*. *Again* resists appearing in the scope of *continuously*:

(95) a. Carlos [again [continuously [relied on his dad]]].
b. *Carlos [continuously [again [relied on his dad]]].

(96a,b) thus shows that preverbal adverbs are attached higher than adverbs sandwiched between the verb and a PP complement (prediction (iii)). (96c,d) shows that adverbs sandwiched between the verb and a PP complement are attached lower than adverbs that follow a PP complement (prediction (iv)). (96e) shows that preverbal adverbs can c-command adverbs that follow a PP complement; (96f) shows that the reverse c-command relation is also possible (prediction (v)).

(96) a. Carlos again relied continuously on his dad.
b. *Carlos continuously relied again on his dad.
c. Carlos relied continuously on his dad again.
d. ??Carlos relied again on his dad continuously.
e. Carlos again relied on his dad continuously.
f. Carlos continuously relied on his dad again.

The predictions can also be tested using pairs of so-called reversable adverbs:

(97) a. Ava repeatedly knocked accidentally on the door. (*repeatedly* > *accidentally*)
b. Ava accidentally knocked twice on the door. (*accidentally* > *twice*)
c. Ava knocked accidentally on the door twice. (*twice* > *accidentally*)
d. Ava knocked twice on the door accidentally. (*accidentally* > *twice*)
e. Ava repeatedly knocked on the door accidentally. (ambiguous)
f. Ava accidentally knocked on the door twice. (ambiguous)

Finally, since temporal adverbs are attached higher than manner adverbs at the very least as a matter of preference (Jackendoff 1972, Cinque 1999 and Ernst 2002), the examples in (98) confirm the three predictions under scrutiny.

(98) a. Ava yesterday talked softly to Carlos.
b. *Ava softly talked yesterday to Carlos.
c. Ava talked softly to Carlos yesterday.
d. *Ava talked yesterday to Carlos softly.
e. Ava yesterday talked to Carlos softly.
f. Ava softly talked to Carlos yesterday.

Many speakers find (98b) and (98d) degraded even if *softly* is omitted, suggesting attaching a time adverb within the verb's selection domain leads to reduced grammaticality. That in itself confirms that adverbs between the verb and a PP complement are attached lower than adverbs that appear preverbally or following a PP complement.

(99) a. Ava talked softly to Carlos.
b. ??Ava talked yesterday to Carlos.

Neeleman and Payne (2020) show that PPs can be attached higher than temporal adverbs more easily in certain circumstances, e.g. when they function as continuing topics in Lambrecht's (1994:132) sense. But even in these circumstances, the orders in (98b,d) remain unacceptable, in line with predictions (iii) and (iv):

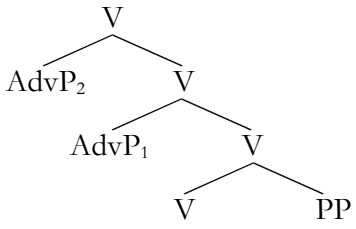
(100) Ava talked yesterday to Carlos, and he told her some news.

(101) a. *Ava softly talked yesterday to Carlos, and he told her some news.
b. *Ava talked yesterday to Carlos softly, and he told her some news.

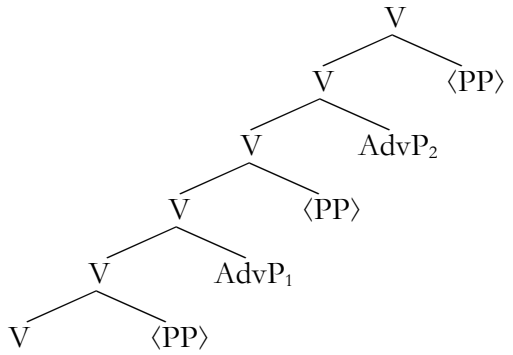
There are three further predictions to consider, which concern adjacent adverbs:

vi. If two adverbs appear preverbally, the leftmost adverb takes scope over the rightmost adverb, (as in (102a)).
vii. If two adverbs appear sandwiched between the verb and a PP complement, the rightmost adverb takes scope over the leftmost adverb (as in (102b) with the PP in the highest position).
viii. If two adverbs follow a PP complement the rightmost adverb takes scope over the leftmost adverb (as in (102b) with the PP in the lowest position).

(102) a.

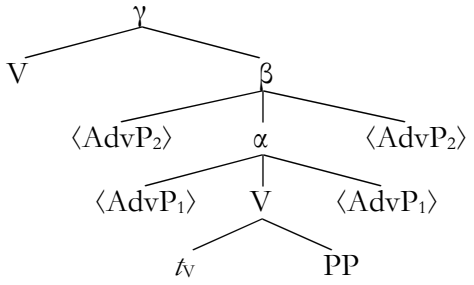


b.



Prediction (vi) is uncontroversially correct. But this not true of prediction (vii). Pesetsky (1989) and Johnson (1991) argue that the V-AdvP-PP order can be derived by (non-case-driven) verb raising, as well as by attaching the PP higher than the adverb:

(103)



The verb raising analysis (combined with the option of extraposition) shares predictions (i), (iii), (v), (vi) and (viii) with the current proposal. However, it diverges on the remaining three predictions. It predicts that…

ii'. An adverbial sandwiched between the verb and a PP complement may c-command the PP complement, or be c-commanded by it if the PP complement is extraposed.
iv'. An adverb sandwiched between the verb and a PP complement may c-command or be c-commanded by an adverb that follows the PP complement.
vii'. If two adverbs appear sandwiched between the verb and a PP complement, either the rightmost adverb takes scope over the leftmost adverb, or the leftmost adverb takes scope over the rightmost adverb.

Predictions (ii') and (iv') are incorrect. I will now look in more detail at predictions (vii/vii') and (viii). For pairs of temporal and manner adverbs sandwiched between the verb and a PP complement, only ascending orders are allowed, in line with prediction (vii). The same restriction holds of such pairs if placed sentence-finally, in line with prediction (viii):

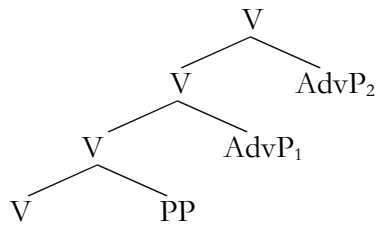
(104) a. Ava talked ⟨softly⟩ yesterday ⟨*softly⟩ to Carlos, and he told her some news.
b. Ava talked to Carlos ⟨softly⟩ yesterday ⟨*softly⟩, and he told her some news.

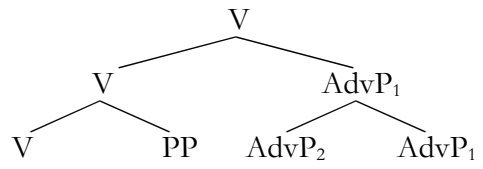
For pairs of low adverbs, the data are incompatible with either proposal. C-command between the adverbs can go either way when they are adjacent (contra predictions (vii) and (viii)) (see also Bobaljik 2017):

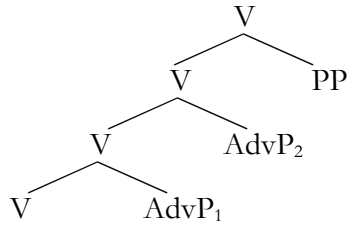
(105) a. Carlos knocked ⟨continuously⟩ again ⟨continuously⟩ on the door.
b. Carlos knocked on the door ⟨continuously⟩ again ⟨continuously⟩.

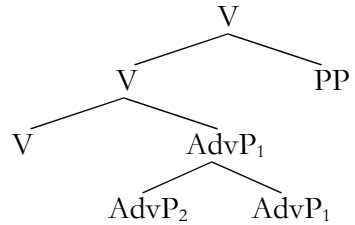
(106) a. Carlos knocked accidentally twice on the door. (ambiguous)
b. Carlos knocked on the door accidentally twice. (ambiguous)

Rohrbacher (1994), Ackema and Neeleman (2002) and Neeleman and Payne (2020) argue that some adverbs can be modified by certain other adverbs (with the first taking scope over the second). This fixes the problem for the verb-in-situ approach:

(107) a.


b.


c.


d.


Clefts show that (relevant) low adverbs can cluster, while temporal and manner adverbs cannot. It follows that predictions (vii) and (viii) hold for pairs of temporal and manner adverbials, but not for *again continuously* and *accidentally twice*.

(108) a. *It was yesterday SOFTLY that Ava talked to Carlos.
b. It was again CONTINUOUSLY that Carlos knocked on the door.
c. It was accidentally TWICE that Carlos knocked on the door.

Furthermore, adverb clustering has interpretative effects that are present in structures that go against predictions (vii) and (viii).

Again triggers a presupposition whose content depends on the material it c-commands. This explains why (109a) is infelicitous in the context given. If *again* is attached to *continuously*, it only triggers the presupposition that there was a previous continuous event, explaining the felicity of (109b,c).

(109) [After Carlos arrived at the house, he rattled the window continuously and then…]
  a. #He again knocked ⟨continuously⟩ on the door ⟨continuously⟩.
  b. It was again continuously that he knocked on the door.
  c. He knocked ⟨on the door⟩ again continuously ⟨on the door⟩.

Next consider *accidentally twice*. *Accidentally* is an adverbial that can associate with a focused constituent in its c-command domain (see Williams 2014):

(110) a. Carlos accidentally knocked on the door TWICE. (111a); *(111b)
  b. Carlos accidentally knocked on the DOOR twice. *(111a); (111b)

(111) a. (i) John gave two knocks on the door; (ii) ∃*n*, *n* and alternative to 2, that John intended to give *n* knocks on the door.
  b. (i) John gave two knocks on the door; ∃*a*, *a* an alternative to the door, such that John intended to give two knocks on *a*.

As association with focus is subject to c-command, *accidentally* in the adverbial cluster *accidentally twice* must associate with *twice*. Thus, (112a) does not allow the interpretation in (111b). Similarly, (112b) does not allow an interpretation in which *accidentally* takes scope over *twice* but associates with *the door*:

(112) a. It was accidentally TWICE that Carlos knocked on the door. (111a); *(111b)
  b. Carlos knocked accidentally twice on the door. (111a); *(111b)

Thus, adverbial clustering reconciles the verb-in-situ approach with counterexamples to predictions (vii) and (viii). The verb-raising analysis cannot use adverbial clustering to explain counterexamples to prediction (viii), because this would remove the strongest evidence in support of verb movement.

A few comments are in order about adverb clustering in context other than the ones discussed above. To begin with, one would expect adverb clustering elsewhere, and indeed it is found sentence-finally in examples like the following:

(113) a. [Carlos agreed to paint two doors once, but he made a mistake.]
  He painted the first door [accidentally twice].
  b. [Ava played the harpsichord continuously for an hour. We thought that was it, but then…]
  She played the recorder [again continuously] for forty minutes.

This does not mean that earlier claims are undermined. For example, I used the examples in (114) as confirmation of the prediction that adverbs within the verb's selection domain come in ascending order. This prediction stands as the relevant adverbs to not cluster.

(114) a. Ava treated Carlos ⟨*twice⟩ <u>badly</u> ⟨twice⟩.
  b. Ava dressed her children ⟨*twice⟩ <u>elegantly</u> ⟨twice⟩. (quality)
  c. John fired two employees ⟨*just now⟩ <u>again</u> ⟨just now⟩. (*2 > again*)

(115) a. *It was [twice badly] that Ava treated Carlos.
  b. *It was [twice elegantly] that Ava dressed her children.
  c. *It was [just now again] that John fired two employees.

Similarly, I claimed in section 1 that the obligatory descending order of adverbs sandwiched between verb and object in Czech, provided evidence for verb raising from an underlying position adjacent to the object. This argument is not affected by the option of adverb clustering. If a given pair of adverbs sandwiched between verb and object can only come in the descending order, as in (116), the structure must be derived by verb movement. Suppose there were no verb movement and adverbs were instead allowed to attach between the base positions of verb and object. Then, the descending order could result from clustering, but the ascending order would be ruled in as well, as the adverbs could attach to the clausal spine separately.

(116) a. V $AdvP_2$ $AdvP_1$ DP
  b. *V $AdvP_1$ $AdvP_2$ DP

Conversely, if the two adverbs can only come in ascending order, as in (117), the structure qualifies as a genuine counterexample to the CFC. Suppose the verb were generated in a position adjacent to the object and subsequently moved across the adverbs. Then the ascending order could be explained as resulting from

clustering (with the right-most adverb modifying the leftmost one). However, the descending order should be grammatical as well, in case the adverbs do not cluster.

(117) a. *V $AdvP_2$ $AdvP_1$ DP
b. V $AdvP_1$ $AdvP_2$ DP

The only case in which no conclusions can be drawn without further research is when there is no fixed order of adverbs. As this is not true in Czech, there remains a strong case for verb movement in this language.

5.4.4. AcI

I finally consider the distribution of adverbs in sentences with an infinitival complement. The CFC predicts that in AcI constructions, neither adverbs belonging to the matrix clause nor adverbs belonging to the embedded clause can appear between the verb and the embedded subject. This explains the contrasts between (118a) and (118b) and between (119a) and (119b).

(118) a. I ⟨confidently⟩ expected ⟨confidently⟩ that [Marcia would leave].
b. I ⟨confidently⟩ expected ⟨*confidently⟩ [Marcia to leave].

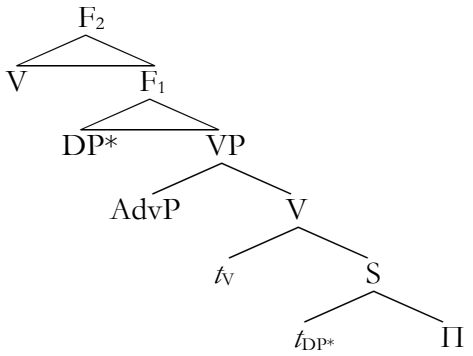
(119) a. I expect that [⟨tomorrow⟩ Marcia will leave ⟨tomorrow⟩].
b. I expect [⟨*tomorrow⟩ Marcia to leave ⟨tomorrow⟩].

These data support the overall approach adopted here. However, as is well known, English permits matrix adverbs to intervene between the embedded subject and predicate of an AcI construction:

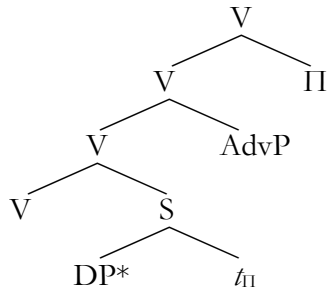
(120) John believed Mary sincerely to be the winner.

This has widely been construed as evidence for raising to object (RtO; Postal 1974, Johnson 1991, Lasnik 1999). However, it could also result from ECM, combined with extraposition of the embedded predicate (ECM+X; Pesetsky 1982, Massam 1985, Neeleman 1994).

(121) a.



b.



Neeleman and Payne (2021b) present evidence in favour of ECM+X. The first argument is based on Barss's generalization. Existentials reconstruct in A-chains (May 1977; Hornstein 1995; Fox 1999; Lebeaux 2009):

(122) a. [Some young lady]$_1$ [seems (to me) [t_1 to be likely t_1 to dance with every senator]]. (∃>∀; ∀>∃)
b. [Some young lady]$_1$ [seems to herself$_1$ [t_1 to be likely t_1 to dance with every senator]]. (*∀>∃)
c. Mary$_1$ [seems to some young lady [t_1 to be likely t_1 to dance with every senator]]. (*∀>∃)

Barss (1986) and Sauerland and Elbourne (2002) observe that reconstruction is blocked if the AP in which the existential originates moves to a position c-commanding the existential (i.e. reconstruction requires surface c-command of the reconstruction site):

(123) [How likely t_1 to dance with every senator]$_2$ does [some young lady]$_1$ [seem [t_1 to be t_2]]? (*∀>∃)

The RtO analysis predicts that scope will be unaffected by adverbial intervention, but the ECM+X analysis predicts Barss's Generalization effects. These are indeed found:

(124) a. John sincerely believed some young lady to be likely to dance with every senator.
(∃>∀; ∀>∃)

b. John believed some young lady sincerely to be likely to dance with every senator.
(*∀>∃)

The pattern is consistent. All the (a) examples below are ambiguous (∃>∀; ∀>∃); none of the (b) examples allows the existential to be interpreted in the scope of the universal (∃>∀; *∀>∃).

(125) a. The spokeswoman reluctantly confirmed some battlecruiser to appear to have bombarded every coastal town.
b. The spokeswoman confirmed some battlecruiser reluctantly to appear to have bombarded every coastal town.

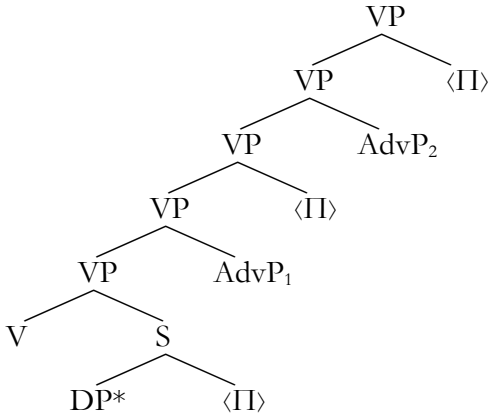
(126) a. The engineer quite easily proved some metal part to be certain to fail in every T-39 aircraft.
b. The engineer proved some metal part quite easily to be certain to fail in every T-39 aircraft.

(127) a. Jeremy privately admitted some politician to be likely to be involved in every current scandal.
b. Jeremy admitted some politician privately to be likely to be involved in every current scandal.

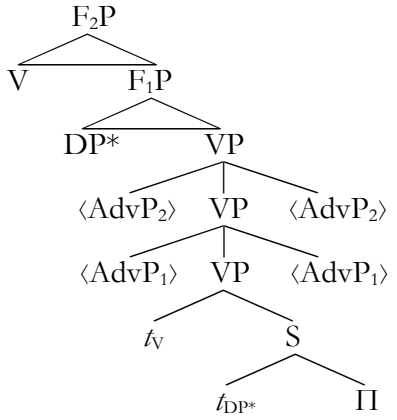
(128) a. Tara incorrectly assumed some doctor to be likely to visit every patient by noon.
b. Tara assumed some doctor incorrectly to be likely to visit every patient by noon.

The second argument is based on the order of adverbs in AcI constructions. The ECM+X analysis predicts ascending order throughout. The RtO analysis predicts descending order when two adverbs intervene, variable order when the embedded predicate appears between the two adverbs and ascending order when both adverbs appear sentence-finally:

(129) a.



b.



The data support the extraposition analysis. Neeleman and Payne provide experimental evidence for the contrasts given below. (The examples in (131) were presented in a context that does not favor adverb clustering, but there is some evidence that it still played a role, as *again-continuously* was judged better when the two adverbs are adjacent.)

(130) a. John believed Mary ⟨sincerely⟩ yesterday ⟨*sincerely⟩ to be six feet tall.
b. John believed Mary sincerely to be six feet tall yesterday.
c. *John believed Mary yesterday to be six feet tall sincerely.
d. John believed Mary to be six feet tall ⟨sincerely⟩ yesterday ⟨*sincerely⟩.

(131) [During their first tour of duty, John continuously expected Bill to die, but this never happened. During their second tour of duty …]
a. John expected Bill ⟨continuously⟩ again ⟨*continuously⟩ to die
b. John expected Bill continuously to die again

c. *John expected Bill again to die continuously
d. John expected Bill to die ⟨continuously⟩ again ⟨*continuously⟩.

In sum, the distribution of adverbs in AcI constructions supports the proposed account of compactness, rather than posing a problem for it.

5. Conclusion

I have argued that the association between the XV/VX parameter on the one hand and adverbial intervention/compactness on the other can be understood in terms of two linear factors: the XV/VX parameter itself (which is taken to be a linearization parameter), and the CFC. The distribution of FQs was accounted for by a further linearization constraint that required placement between an argument and a predicate.

There is circumstantial evidence that the linear constraints in question are part of the syntax: (i) they drive VP-shell formation, a syntactic process; (ii) linearization with respect to *traces* matters, even though traces have no phonological realization.

If so, the proposal is at odds with the claim in Chomsky 1995, 2013, Berwick et al. 2011 and Chomsky et al. 2019 that linear order is a reflex of the sensory-motor system and plays no role in syntax or semantics. These authors provide strong evidence that some syntactic rules are structure-dependent, but of course this does not imply that no syntactic rule refers to linear order.

Selected references

Abels, K. 2003. *Successive Cyclicity, Anti-locality, and Adposition Stranding*. PhD dissertation, University of Connecticut.

Abels, K. 2016. The Fundamental Left–Right Asymmetry in the Germanic Verb Cluster. *JCGL* 19: 179–220.

Abels, K., and A. Neeleman. 2012. Linear asymmetries and the LCA. *Syntax* 15: 25–74.

Abrusán, Márta. 2014. *Weak island semantics*. OUP.

Altshuler, Daniel, and Robert Truswell. 2022. *Coordination and the syntax-discourse interface*. OUP.

Ambridge, Ben; Julian M. Pine; and Elena V.M. Lieven 2014. Child language acquisition: Why Universal Grammar doesn't help. *Language* 90: 53–90.

Aoun, J. 1982. On the logical nature of the binding principles: Quantifier lowering, double raising of *there* and the notion empty element. In *Proceedings of NELS 12* (16–36). GLSA, University of Massachusetts, Amherst.

Baltin, M. 1978. *Toward a Theory of Movement Rules*. PhD dissertation, MIT.

Baltin, M. 1982. A Landing Site Theory of Movement Rules. *LI* 13: 1–38.

Baltin, M. 1995. Floating Quantifiers, PRO, and Predication. *LI* 26: 199–248.

Barbiers, S. 1995. *The Syntax of Interpretation*. PhD dissertation, University of Leiden.

Barker, C. 2012. Quantificational binding does not require c-command. *LI* 43: 614–633.

Barss, A. 1986. *Chains and Anaphoric Dependence*. PhD dissertation. MIT.

Bayer, S. 1996. The coordination of unlike categories. *Language* 72: 579–616.

Belletti, A. 1982. On the Anaphoric Status of the Reciprocal Construction in Italian. *The Linguistic Review* 2: 101–137.

Bennis, Hans. 1986. *Gaps and Dummies*. Berlin: Foris.

Berwick, R. et al. 2011. Poverty of the Stimulus Revisited. *Cognitive Science* 35: 1207–1242.

Bobaljik, J. 1995. *Morphosyntax: The Syntax of Verbal Inflection*. PhD dissertation, MIT.

Bobaljik, J. 1999. Adverbs: The hierarchy paradox. *Glot International* 4.9/10.

Bobaljik, J., and Brown, S. 1997. Interarboreal operations: Head movement and the extension requirement. *LI* 28: 345–356.

Borgonovo, C. 1994. *The parametric syntax of gerunds*. Waltham, MA: Brandeis University dissertation.

Borgonovo, C., and A. Neeleman. 2000. Transparent adjuncts. *Canadian Journal of Linguistics* 45: 199–224.

Borsley, R. 1994. In defense of coordinate structures. *Linguistic Analysis* 24: 218–245.

Borsley, R. 2005. Against ConjP. *Lingua* 115: 461–482.

Bošković, Ž. 2020. On the coordinate structure constraint, across-the-board-movement, phases, and labeling. In. *Recent developments in phase theory,* ed. by J. van Craenenbroeck, C. Pots and T. Temmerman (133–182). Berlin: De Gruyter Mouton.

Bošković, Ž., and S. Franks. 2000. Across-the-board movement and LF. *Syntax* 3: 107–128.

Bowers, J. 2001. Predication. *The handbook of contemporary syntactic theory*, ed. by M. Baltin and C. Collins (299–333). Oxford: Blackwell Publishers.

Boyce, V., R. Futrell and R. Levy. 2020. Maze made easy: Better and easier measurement of incremental processing difficulty. *Journal of Memory and Language* 111: 104082.

Bresnan, J. 1974. The position of certain clause-particles in phrase structure. *LI* 5: 614–619.

Brown, Jessica. 2017. *Heads and Adjuncts: An Experimental Study of Subextraction from Participials and Coordination in English, German and Norwegian*. PhD dissertation, University of Cambridge.

Bruening, B. 2001. QR obeys Superiority: ACD and frozen scope. *LI* 32: 233–273.

Bruening, B., and E. Al Khalaf. 2020. Category mismatches in coordination revisited. *LI* 51: 1–36.

Büring, D. 2024. Flexible Semantics. Ms. University of Vienna.

Büring, D., and K. Hartmann. 2015. Semantic coordination without syntactic coordinators. In *Structures in the Mind: Essays on Language, Music, and Cognition in Honor of Ray Jackendoff*, ed. by I. Toivonen et al. (41–62). Cambridge, MA: MIT Press.

Carden, Guy, and David Pesetsky. 1977. Double-verb constructions, markedness, and a fake co-ordination. *Papers from the 13th regional meeting of the Chicago Linguistic Society*, ed. by W.A. Beach et al. (82–92). Chicago: CLS.

Carrier, Jill, and Janet H. Randall 1992. The argument structure and syntactic structure of resultatives. *LI* 23: 173–234.

Chaves, R. 2007. *Coordinate Structures and Constraint-Based Syntax-Semantics Processing*. PhD dissertation, University of Lisbon.

Chaves, R. 2012. On the grammar of extraction and coordination. *NLLT* 30: 465–512.

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. 1973. Conditions on transformations. In *A festschrift for Morris Halle*, ed. by S. Anderson and P. Kiparsky (232–286). New York: Holt, Rinehart, and Winston.

Chomsky, N. 1977. *Essays on Form and Interpretation*. New York: North-Holland.

Chomsky, N. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, N. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. 2000. Minimalist inquiries: The framework. In *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. by R. Martin et al. Cambridge, MA: MIT Press. 89–155.

Chomsky, N. 2013. Problems of Projection. *Lingua* 130: 33–49.

Chomsky, N., et al. 2019. Generative Grammar and the Faculty of Language: Insights, Questions, and Challenges. *Catalan Journal of Linguistics* Special Issue: 229–261.
Cinque, G. 1990. *Types of A-Dependencies*. Cambridge, MA: MIT Press.
Cinque, G. 1999. *Adverbs and Functional Heads*. OUP.
Cinque, G. 2005. Deriving Greenberg's Universal 20 and its exceptions. *LI* 36: 315–332.
Cinque, G. 2009. The Fundamental Left-Right Asymmetry of Natural Languages. In *Universals of Language Today*, ed. by S. Scalise et al. (165–184). Dordrecht: Springer.
Cinque, G. 2014. On the movement account of Greenberg's Universal 20: Refinements and replies. Ms. University of Venice.
Cinque, G. 2023. Again on the order of demonstrative, numeral, adjective, and noun. Ms. University of Venice.
Collins, C. 1997. *Local Economy*. Cambridge, MA: MIT Press.
Corbett, G. 2006. *Agreement*. CUP.
Cormack, A., and R. Breheny. 1994. Projections for functional categories. *UCL Working Papers in Linguistics* 6: 35–61.
Culbertson, J., P. Smolensky and G. Legendre. 2012. Learning Biases Predict a Word Order Universal. *Cognition* 122: 306–329.
Dalrymple, M., and A. Kehler. 1995. On the constraints imposed by *respectively*. *LI* 26: 531–536.
Deane, P. 1991. Limits to attention: A cognitive theory of island phenomena. *Cognitive Linguistics* 2: 1–63.
Demonte, V. 1987. Remarks on secondary predicates: C-command, extraction, and reanalysis. *The Linguistic Review* 6: 1–39.
Den Dikken, M. 1995. *Particles; On the Syntax of Verb-Particle, Triadic and Causative Constructions*. OUP.
Den Dikken, M. 2006. *Relators and Linkers; The Syntax of Predication, Predicate Inversion, and Copulas*. Cambridge, MA: MIT Press.
De Vos, M. 2005. *The Syntax of Verbal Pseudo-Coordination in English and Afrikaans*. PhD dissertation, University of Leiden.
De Vos, M. 2007. *And* as an aspectual connective in the event structure of pseudo-coordinative constructions. In *Connectives as discourse landmarks*, ed. by A. Celle and R. Huart (49–70). Amsterdam: John Benjamins Publishing.
De Vos, M. 2009. An argument for VP coordination: Scene-setting coordination and weak islands. *Southern African Linguistics and Applied Language Studies* 27: 1–23.
De Vries, M. 2005. Coordination and syntactic hierarchy. *Studia Linguistica* 59.1.83–105.
De Vries, M. 2017. Across-the-board phenomena. In *The Wiley Blackwell Companion to Syntax*, Second Edition, ed. by Martin Everaert and Henk van Riemsdijk (1–31). Hoboken: John Wiley and Sons.
Dik, S. 1968. *Coordination: Its Implications for the Theory of General Linguistics*. Amsterdam: North-Holland.
Dik, S. 1983. Two constraints on relators and what they can do for us. In *Advances in Functional Grammar*, ed. by S. Dik (267–298). Dordrecht: Foris.
Dik, S. 1997. *The Theory of Functional Grammar; Part 1: The Structure of the Clause* (second edition). Berlin: Mouton de Gruyter.
Doetjes, J. 1997. *Quantifiers and Selection*. PhD dissertation, University of Leiden.
Dowty, D. and B. Brody. 1984. A Semantic Analysis of 'Floated' Quantifiers in Transformationless Grammar. In *Proceedings of WCCFL 3*, ed. by M. Cobler et al. (75–90). Stanford: Stanford Linguistics Association.
Dryer, M. 1991. SVO languages and the OV: VO typology. *JoL* 27: 443–482.
Dryer, M. 2018. On the order of demonstrative, numeral, adjective and noun. *Language* 94: 798–833.
Emonds, J. 1978. The Verbal Complex V'-V in French. *LI* 9: 49–77
Ernst, T. 2002. *The Syntax of Adjuncts*. CUP.
Fábregas, A., and Á. Jiménez-Fernández. 2016. Extraction from gerunds and the internal syntax of verbs. *Linguistics* 54:1307–1354.
Fanselow, G. 1990. Scrambling as NP-Movement. In *Scrambling and Barriers*, ed. by G. Grewendorf and W. Sternefeld (eds.) (113–140). Amsterdam: John Benjamins.
Fodor, J.D., and A. Inoue. 1994. The diagnosis and cure of garden paths. *Journal of Psycholinguistic Research* 23: 407–434.
Fujita, H. 2021. On the parsing of garden-path sentences. *Language, Cognition and Neuroscience* 36: 1234–1245.
Forster, K., C. Guerrera and L. Elliot. 2009. The maze task: Measuring forced incremental sentence processing time. *Behavior Research Methods* 41: 163–171.
Fox, D. 1999. Reconstruction, binding theory, and the interpretation of chains. *LI* 30: 157–196.
Fox, Danny. 2000. *Economy and Semantic Interpretation*. Cambridge, MA: MIT Press.
Georgi, D., and G. Müller. 2010. Noun-phrase structure by reprojection. *Syntax* 13: 1–36.
Goldberg, A. 2006. *Constructions at Work: The Nature of Generalization in Language*. OUP.
Goodall, G. 1987. *Parallel Structures in Syntax: Coordination, Causatives, and Re-Structuring*. CUP.
Greenberg, J. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of language*, ed. by J. Greenberg (73–113). Cambridge, MA: MIT Press.
Grimshaw, J. 2005. *Words and structure*. Stanford, CA: CSLI.
Grosu, A. 1973. On the nonunitary nature of the coordinate structure constraint. *LI* 4: 88–92.
Haider, H. 2014. *Symmetry Breaking in Syntax*. CUP.
Hartmann, J. 2008. *Expletives in Existentials: English* there *and German* da. Doctoral thesis. Tilburg University.

Hartmann, K. 2000. *Right Node Raising and Gapping: Interface Conditions on Prosodic Deletion*. Philadelphia: John Benjamins Publishing.
Haspelmath, M. 2007. Coordination. In: *Language Typology and Linguistic Description; volume II: Complex Constructions*, ed. by T. Shopen (1–51). CUP.
Haspelmath, M. 2013. Nominal and verbal conjunction. In: *The World Atlas of Language Structures Online*, ed. by M. Dryer and M. Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. (http://wals.info/chapter/64, accessed 10 August 2020).
Hawkins, J. 1990. A parsing theory of word order universals. *LI* 21: 223–261.
Hawkins, J. 2008. An Asymmetry between VO and OV Languages; The Ordering of Obliques. In *Case and grammatical Relations*, ed. by G. Corbett and M. Noonan (eds.) (167–190). Amsterdam: John Benjamins.
Hawkins, J. 2009. Language universals and the performance-grammar correspondence hypothesis. In *Language Universals*, ed. by M. Christiansen, C. Collins and S. Edelmann (54–78). OUP.
Hazout, I. 2004. The syntax of existential constructions. *LI* 35.3.393–430.
Heim, I., and A. Kratzer. 1998. *Semantics in Generative Grammar*. Malden: Blackwell.
Higginbotham, J. 1985. On semantics. *LI* 16.4.547–593.
Hoekstra, T. 1984. *Transitivity: Grammatical Relations in Government-Binding theory*. Dordrecht: Foris Publications.
Hofmeister, Philip, and Ivan A. Sag. 2010. Cognitive constraints and island effects. *Language* 86: 366–415.
Hornstein, N. 1995. *Logical Form; From GB to Minimalism*. Cambridge: Blackwell.
Hornstein, N., and M. Polinsky. 2010. *The Movement Theory of Control*. Amsterdam: John Benjamins.
Hornstein, N., and A. Weinberg. 1990. The necessity of LF. *The Linguistic Review* 7: 129–167.
Huang, C.-T. James. 1982. *Logical Relations in Chinese and the Theory of Grammar*. Cambridge, MA: MIT dissertation.
Huybregts, R., and H. van Riemsdijk. 1985. Parasitic gaps and ATB. In *Proceedings of NELS 15* (168–187). University of Massachusetts, Amherst: GSLA.
Jackendoff, R. 1996. *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.
Johannessen, J.B. 1998. *Coordination*. OUP.
Johnson, K. 1991. Object Positions. *NLLT* 9: 577–636.
Johnson, K. 2004. In search of the English middle field. Ms., University of Massachusetts at Amherst.
Just, M., P. Carpenter and J. Woolley. 1982. Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General* 111: 228–238.
Kamp, H., and B. Partee. 1995. Prototype theory and compositionality. *Cognition* 57: 129–191.
Kayne, R. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris Publications.
Kayne, R. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
Kehler, A. 2002. *Coherence, Reference, and the Theory of Grammar*. Stanford: CSLI Publications.
Kentner, G., and C. Féry. 2013. A new approach to prosodic grouping. *The Linguistic Review* 30: 277–311.
Keyser, S.J. 1968. Review of Sven Jacobson, *Adverbial Positions in English*. *Language* 44: 357–374
Kim, N., L. Brehm, P. Sturt, and M. Yoshida. 2020. How long can you hold the filler: maintenance and retrieval. *Language, Cognition and Neuroscience* 35: 17–42.
Kirby, S. 1999. *Function, selection, and innateness: The emergence of language universals*. OUP.
Koeneman, O. 2000. *The Flexible Nature of Verb Movement*. Doctoral thesis. Utrecht University.
Koster, J. 1974. Het werkwoord als spiegelcentrum. *Spektator* 3: 601–618.
Kluender, R. 1991. *Cognitive Constraints on Variables in Syntax*. PhD dissertation, University of California at San Diego.
Kubota, Y., and J. Lee. 2015. The coordinate structure constraint as a discourse-oriented principle: Further evidence from Japanese and Korean. *Language* 91: 642–675.
Kush, Dave; Terje Lohndal; and Jon Sprouse. 2018. Investigating variation in island effects: A case study of Norwegian wh-extraction. NLLT 36: 743–779.
Lakoff, G. 1986. Frame semantic control of the coordinate structure constraint. In *Proceedings of the 22nd regional meeting of the Chicago Linguistic Society, vol. 2, Papers from the parasession on pragmatics and grammatical theory*, ed. by A.M. Farley, P.T. Farley and K.-E. McCullough (152–167). Chicago: CLS.
Landau, I. 2003. Movement out of control. *LI* 34: 471–498.
Landau, I., and J. Bobaljik. 2009. Icelandic control is not A-movement: The case from case. *LI* 40: 113–154.
Larson, Richard. 1988. On the Double-Object Construction. *LI* 19: 335–391.
Lasnik, H. 1999. Chains of arguments. In *Working Minimalism*, ed. by S. Epstein and N. Hornstein (189–215). Cambridge, MA: MIT Press.
Lasnik, H., and M. Saito 1984. On the nature of Proper Government. *LI* 15: 235–289.
Lasnik, H., and M. Saito 1992. *Move-Alpha: Conditions on its Application and Output*. Cambridge, MA: MIT Press.
Lebeaux, D. 2009. *Where does the Binding Theory Apply?* Cambridge, MA: MIT Press.
Lechner, W. 1998. Two kinds of reconstructions. *Studia Linguistica* 52: 276–310.
Lechner, W. 2013. Diagnosing covert movement: The Duke of York and reconstruction. In *Diagnosing Syntax*, ed. by L. Cheng and N. Corver (158–189). OUP.
Lechner, W. 2019. A calculus for reconstruction and anti-reconstruction. In *Reconstruction Effects in Relative Clauses*, ed. by M. Krifka and M. Schenner (113–144). Berlin: De Gruyter.
Levelt, W. 1989. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levin, B., and M. Rappaport Hovav. 1995. *Unaccusativity: At the syntax-lexical semantics interface*. Cambridge, MA: MIT Press.
Maienborn, C. 2001. On the position and interpretation of locative modifiers. *Natural Language Semantics* 9: 191–240.
Maling, J. 1976. Notes on Quantifier-Postposing. *LI* 7: 708–718.
Marušič, F., A. Nevins and B. Badecker. 2015. The grammars of conjunction agreement in Slovenian. *Syntax* 18.1.39–77.
Massam, D. 1985. *Case Theory and the Projection Principle*. PhD dissertation, MIT.
May, R. 1977. *The Grammar of Quantification*. PhD dissertation, MIT, Cambridge, MA.
May, R. 1985. *Logical form: Its structure and derivation*. Cambridge, MA: MIT Press.
Moro, Andrea. 1997. The Raising of Predicates: Predicative Noun Phrases and the Theory of Clause Structure. CUP.
Muadz, H. 1991. *Coordinate Structures: A Planar Representation*. Doctoral thesis. University of Arizona.
Munn, A. 1987. Coordinate structure and X-bar theory. *McGill Working Papers in Linguistics* 4: 121–140.
Munn, A. 1992. A null operator analysis of ATB gaps. *The Linguistic Review* 9: 1–26.
Munn, A. 1993. Topics in the Syntax and Semantics of Coordinate Structures. PhD dissertation, University of Maryland.
McNulty, E., 1988. *The Syntax of Adjunct Predicates*. PhD dissertation, University of Connecticut.
Na, Y., and G.J. Huck. 1992. *On Extracting from Asymmetrical Structures*. In *The Joy of Grammar: A Festschrift in Honor of James D. McCawley*, ed. by D. Brentari et al. (251–274). Amsterdam: John Benjamins.
Nakajima, H. 1990. Secondary predication. *The Linguistic Review* 7: 275–309.
Neeleman, A. 1994. *Complex Predicates*. PhD dissertation, Utrecht University.
Neeleman, A. 2017. PP-over-V meets Universal 20. *JCGL* 20: 3–47.
Neeleman, A. and H. van de Koot. 2020. The Non-Existence of Sub-Lexical Scope. In *Linguistic Variation*, ed. by G. Bellucci et al. (501–530). Berlin: De Gruyter.
Neeleman, A. and H. van de Koot. 2022. The interpretation and distribution of temporal focus particles. *NLLT* 40: 793–835
Neeleman, A, and F. Weerman. 1999. *Flexible Syntax*. Dordrecht: Kluwer.
Nevins, A., and P. Weisser. 2019. Closest conjunct agreement. *Annual Review of Linguistics* 5: 219–241.
Nilsen, Ø. 2003. *Eliminating Positions*. PhD dissertation, University of Utrecht
Parsons, T. 1990. *Events in the Semantics of English: A Study in Subatomic Semantics*. Cambridge, MA: MIT Press.
Patejuk, A. and Przepiórkowski, A. 2023. Category mismatches in coordination vindicated. *LI* 54: 326–349.
Paul, W. 2012. Why Chinese DE is not like French DE: A critical analysis of the predicational approach to nominal modification. *Studies in Chinese Linguistics* 33: 183–210.
Payne, J. 1985. Complex phrases and complex sentences. In: *Language Typology and Linguistic Description; volume II: Complex Constructions*, ed. by T. Shopen. (3–41). CUP.
Pesetsky, D. 1982. *Paths and Categories*. PhD dissertation, MIT.
Pesetsky, D. 1989. Language-particular processes and the Earliness Principle. Ms., MIT.
Philip, J. 2012. *Subordinating and Coordinating Linkers*. Doctoral thesis. University College London.
Philip, J. 2013. (Dis)harmony, the Head-Proximate Filter, and Linkers. *JoL* 49.1.165–213.
Phillips, C., and M. Wagers. 2007. Relating structure and time in linguistics and psycholinguistics. In *The Oxford handbook of psycholinguistics*, ed. by G. Gaskell (739–756). OUP.
Poole, E. and S. Keine. 2024. Not all reconstruction effects are syntactic. *NLLT*.
Postal, Paul. 1998. *Three Investigations of Extraction*. Cambridge, MA: MIT Press.
Prince, A., and P. Smolensky. 2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden: Blackwell.
Progovac, L. 2000. Coordination, c-command, and logophoric N-words. In *Negation and Polarity: Syntactic and Semantic Perspectives*, ed. by L. Horn and Y. Kato (88–114). OUP.
Progovac, L. 2003. Structure for Coordination. In *The Second Glot International State-of-the-Article Book*, ed. by L. Cheng and R. Sybesma (241–288). Berlin: Mouton De Gruyter.
Przepiórkowski, A. 2021. Coordination of unlike grammatical cases. Ms. Polish Academy of Sciences/University of Warsaw.
Pullum, G. and R. Huddleston. 2002. *The Cambridge Grammar of the English Language*. CUP.
Reinhart, T. 1983. *Anaphora and semantic interpretation*. London: Croom Helm.
Reuland, E. 2011. *Anaphora and Language Design*. Cambridge, MA: MIT Press.
Rodman, R. 1976. Scope phenomena, "movement transformations", and relative clauses. In *Montague Grammar*, ed. by B. Partee. New York: Academic Press. 165–176.
Ross, J.R. 1967. *Constraints on Variables in Syntax*. Doctoral thesis. MIT.
Rothstein, S. 1983. *The Syntactic Forms of Predication*. Cambridge, MA: MIT dissertation.
Rothstein, S. 2017. Secondary predication. In *The Wiley Blackwell Companion to Syntax* (second edition), ed. by M. Everaert and H. van. Riemsdijk (209–233). Hoboken, NJ: Wiley Blackwell.
Ruys, E. 1992. *The Scope of Indefinites*. Doctoral thesis. Utrecht University.
Ruys, E. 2010. Expletive selection and CP arguments in Dutch. *JCGL* 13.2.141–178.
Saito, M., and N. Fukui. 1998. Order in Phrase Structure and Movement. *LI* 29: 439–474.

Sag, I.A., G. Gazdar, T. Wasow and S. Weisler. 1985. Coordination and how to distinguish categories. *NLLT* 3: 117–171.
Sauerland, U., and Elbourne, P. 2002. Total reconstruction, PF movement, and derivational order. *LI* 33: 283–319.
Schmerling, S. 1972. Apparent counterexamples to the coordinate structure constraint: A canonical conspiracy. *Studies in the Linguistic Sciences* 2: 91–104.
Sprouse, J., M. Wagers and C. Phillips. 2013. Deriving competing predictions from grammatical approaches and reductionist approaches to island effects. In *Experimental Syntax and Island Effects*, ed. by J. Sprouse (21–41). CUP.
Sprouse, J., I. Caponigro, C. Greco and C. Cecchetto. 2016. Experimental syntax and the variation of island effects in English and Italian. *NLLT* 34: 307–344.
Staub, A., C. Clifton and L. Frazier. 2006. Heavy NP Shift is the parser's last resort: Evidence from eye movements. *Journal of Memory and Language* 54: 389–406.
Sternefeld, Wolfgang. 2001. Syntactic vs. semantic reconstruction. In *Linguistic Form and its Computation*, ed. by C. Rohrer et al. (145–182). Stanford: CSLI.
Stowe, L. 1986. Parsing WH-constructions: Evidence for online gap location. *Language and Cognitive Processes* 1: 227–245.
Stowell, T. 1981. *Origins of Phrase Structure*. Doctoral dissertation, MIT.
Szabolcsi, A. 2006. Strong vs. weak islands. In *The Blackwell companion to syntax*, ed. by M. Everaert and H. van Riemsdijk (vol. 4, 479–531). Malden, MA: Blackwell Publishing.
Truswell, Robert. 2011. *Events, Phrases, and Questions*. OUP.
Van Craenenbroeck, J. 2020. Expletives, locatives and subject doubling. In *Linguistic Variation: Structure and Interpretation*, ed. by L. Franco and P. Lorusso (661–690). Berlin: De Gruyter Mouton.
Van de Koot, H., R. Silva, C. Felser and M. Sato. 2015. Does Dutch A-Scrambling Involve Movement? Evidence from Antecedent Priming. *The Linguistic Review* 32: 739–776.
Wagers, M., and C. Phillips. 2009. Multiple dependencies and the role of the grammar in real-time comprehension. *JoL* 45: 395–433.
Wagner, M. 2010. Prosody and recursion in coordinate structures and beyond. *NLLT* 28: 183–237.
Wasow, T. 1997a. End-weight from the speaker's perspective. *Journal of Psycholinguistic Research* 26: 347–361.
Wasow, T. 1997b. Remarks on grammatical weight. *Language Variation and Change* 9: 81–105.
Weisser, P. 2015. *Derived Coordination: A Minimalist Perspective on Clause Chains, Converbs and Asymmetric Coordination*. Berlin: De Gruyter.
Weisser, P. 2020. On the symmetry of case in conjunction. *Syntax* 23: 42–77.
Wiklund, A.-L. 2007. *The Syntax of Tenselessness: Tense/Mood/Aspect-agreeing Infinitivals*. Berlin: Mouton de Gruyter.
Williams, E. 1980. Predication. *LI* 11: 203–238.
Williams, E. 1986. NP Trace in Theta Theory. *Linguistics and Philosophy* 4: 433–447.
Williams, E. 1994. *Thematic Structure in Syntax*. Cambridge, Mass.: MIT Press.
Williams, E. 2006. Circumstantial Evidence for Dative Shift. In *Organizing Grammar*, ed. by H. Broekhuis et al. (661–668). Berlin: Mouton.
Winter, Y. 2006. Multiple Coordination: Meaning Composition vs. the Syntax-Semantics Interface. Ms. Technion/NIAS.
Wood, J. 2012. Against the Movement Theory of Control: Another Argument from Icelandic. *LI* 43: 322–330.
Wurmbrand, S. 2014. The Merge Condition: A syntactic approach ro selection. In *Minimalism and Beyond: Radicalizing the Interfaces*, ed. by P. Kosta et al. (130–166). Amsterdam: John Benjamins.
Zhang, N. 2010. *Coordination in syntax*. CUP.
Zoerner, E. 1995. *Coordination: The Syntax of andP*. PhD dissertation, University of California, Irvine.
Zwart, J.-W. 2009. Relevance of typology to minimalist inquiry. *Lingua* 119: 1589–1606.
Zwart, J.-W. 1993. *Dutch Syntax*. PhD dissertation, University of Groningen.