28
Syntax and Externalization

Linearization

Ad Neeleman

28.1 Introduction

One of the achievements of generative linguistics is that it has established beyond doubt that both syntactic and interpretive rules are sensitive to syntactic structure. However, there are undeniable linear aspects to the grammar as well, if only because words in spoken language, and to a lesser degree signs in signed languages, must come in a string. This leads to the question of linearization: How are the structural and linear aspects of the grammar connected?

This question subsumes two more detailed issues. The first concerns the mapping rules that connect structure and linear order. The traditional view is that each pair of sister nodes can in principle be linearized independently, so that a structure with three terminals A, B, and C can be mapped onto four strings. If B and C are sisters and A is sister to the node that immediately dominates B and C, the following linearizations are possible: [A [B C]], [[B C] A], [A [C B]], and [[C B] A]. Two other potential linearizations, B-A-C and C-A-B, are ruled out by the *No Tangling Constraint*, which bans crossing branches (see Partee et al. 1993; exceptions are often analyzed in terms of

prosodic inversion (Halpern 1995) or local dislocation (Marantz 1988; Embick and Noyer 2001)).

A highly disruptive proposal by Kayne (1994~~3~~) aimed to limit the linearizations permitted for any given structure to exactly one. Kayne restricts syntactic structures to a template according to which a head can combine with maximally one complement and maximally one specifier, with the specifier preceding the head and the complement following it. So, if in the structures discussed above, B is the head, then the only permitted linearization is [A [B C]]. This hypothesis, known as *Antisymmetry*, has given rise to a range of novel analyses of familiar phenomena, but perhaps more importantly it has placed linear asymmetries (such as contrasts between leftward and rightward movement) firmly on the research agenda.

Antisymmetry constitutes a radical simplification of the mapping procedure. However, that simplification does not come for free, as the word-order variation found in natural languages must be accounted for one way or another. Given that the mapping between structure and order is assumed to be invariant, word-order variation must be accounted for by assuming that languages with different word orders have different structures (or more precisely that the elements pronounced occupy different positions in a structure that may be assumed to be universal, as in the Cartography framework; see Rizzi 1997, Cinque 1999, among many others). This implies that a larger number of movement parameters must be assumed than in the traditional framework, and a concomitant larger number of landing sites. Abels and

Neeleman (2012) show that the number of potential movement parameters in Antisymmetry is equal to the number of potential linearization parameters in more traditional views of phrase structure. In other words, the complexity of linearization in the traditional framework is replaced in Antisymmetry by complexity in the transformational component.

The second issue that research into linearization should provide an answer to is what aspects of the grammar are sensitive to linear order. One view, which goes back decades and has recently been advocated in Berwick and Chomsky (2016), is that linear order is only introduced in the mapping between syntax and phonology. In other words, syntax and semantics are order-free, and hence syntactic and semantic rules cannot refer to the order of the elements they operate on. One argument that Chomsky has given for this position in various publications concerns movement. In English, the derivation of yes/no questions includes fronting of auxiliaries and copulas (if present). This movement must involve the structurally, rather than the linearly, closest verbal head. Thus, moving the main clause copula in (1a) yields a grammatical structure (see (1b)), but moving the copula in the relative clause does not (see (1c)).

(1) a. A man [that is rich] is happy
    b. Is$_1$ a man [that is rich] t_1 happy?
    c. *Is$_1$ a man [that t_1 rich] is happy?

Crucially, no language has been described as having a movement rule that generates outputs like (1c), while there are several languages that have structures like (1b). Chomsky and collaborators present this as an argument for

the claim that linear order is absent in syntax. The argument is persuasive in showing that there are presumably universal structural conditions on movement chains. However, the data do not warrant the conclusion that no aspect of the syntax is sensitive to linear order (see also Willer Gold et al. 2018 and Kayne 2020, §9). After all, the fact that one aspect of head movement is structural does not provide an empirical basis for the inference that all of syntax is structural. Thus, the question of what aspects of grammar are sensitive to linear order remains open.

In this chapter, I explore several phenomena that bear on the two questions raised above: What mapping rules connect structure and order, and which aspects of the grammar are sensitive to linear order? I begin, in section 28.2, by looking at Universal 20, the observation that word order among modifiers in the noun phrase is constant before the noun, but variable after it. There is general agreement that this situation requires a grammar with two components: a mirroring device (either variation in linearization or roll-up movement) and a linear constraint on movement of (constituents containing) the noun, which must be leftward. In section 28.3, I explore limitations on the mirroring device, that is, harmony constraints favoring constancy in the linearization of heads and dependents. In section 28.4, I turn to potential linear constraints on dependencies (including movement of (constituents containing) the lexical head). In section 28.5, I consider adjacency phenomena, that is, situations in which elements must be contiguous in the linear string. Finally, in

section 28.6, I return to the question of which aspects of grammar are sensitive to linear order, considering the data surveyed in previous sections.

28.2 Universal 20 and Its Extensions in Other Domains

Universal 20, as originally proposed by Greenberg (1963), is given in (2). It states that there is an asymmetry in the order of pre- and postnominal modifiers. Relevant prenominal modifiers come in a fixed order, but the order of postnominal modifiers is variable.

> (2)   **Greenberg's Universal 20**
>
> 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.

Universal 20 is best understood to be a generalization over "neutral" or "basic" orders. Thus, there are languages like English in which the neutral order in the noun phrase is Demonstrative–Numeral–Adjective–Noun. Similarly, there are languages like Kîîtharaka and Gungbe (see (3)), which respectively have N-Dem-Num-A or N-A-Num-Dem as their neutral order. However, there are no languages in which A-Num-Dem-N is attested as the neutral order. Some typological evidence to this effect is given in Table 28.1 (where mirror-image orders are placed next to each other and the unattested order is shaded).

(3) a. these   five   empty   bottles

     b. i-kombe   bi-bi   bi-tano   bi-tune   Kîîtharaka (Peter Muriungi, p.c.)

       8-cup   8-this   8-five   8-red

       'these five red cups'

     c. távò   ɖàxó   xóxó   àtɔn   éhè   lɔ   lɛ́   Gungbe (Aboh 2004)

       table   big   old   three   DEM   SPF   PL

       'these three big old tables'

Table 28.1

Four U20 orders with number of attestations (languages, genera). Numbers according to Cinque 2014 precede numbers according to Dryer 2018

Dem Num A N (507, 126) (113, 57)	N A Num Dem (358, 107) (182, 85)
N Dem Num A (69, 24) (8, 6)	A Num Dem N (0,0) (0, 0)

	Cinque 2014	Dryer 2018
Dem Num A N	507, 126	113, 57
N A Num Dem	358, 107	182, 85
N Dem Num A	69, 24	8, 6
A Num Dem N	0, 0	0, 0

The core insight into the analysis of Universal 20 is due to Cinque (1996, 2005, 2014). Cinque shows that the data fall into place once one makes three assumptions. (i) Demonstratives are universally attached higher than numerals and numerals higher than adjectives (according to the hierarchy Dem > Num > A > N). (ii) There is a mirroring device that relates (among other things) the base order Dem-Num-A-N to its mirror image, N-A-Num-Dem (this device could be either roll-up movement or variation in linearization of sister nodes; see below). (iii) (Constituents containing) the noun may move leftward but not rightward. The analysis is schematically represented in Figure 28.1, where "Movement" stands for non-roll-up movement.

Figure 28.1 The analysis of Universal 20

The analysis extends to the full set of attested and unattested orders described in Cinque (2005, 2014) (see Table 28.2). The mirroring device connects the orders in columns I and II, where each grammatical structure has a grammatical counterpart with reversed order. The derivation of the orders in column III crucially involves leftward (non-roll-up) movement of (a constituent containing) N. Hence, none of these orders has a grammatical mirror image. (The remaining shaded orders cannot be generated because the modifiers are attached in violation of the hierarchy Dem > Num > A > N; I cannot demonstrate this here.).

Table 28.2

The full U20 pattern

	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

For example, the order in (Ic) can be generated in accordance with the hierarchy (as in [Dem [[A N] Num]]) and it therefore has a mirrored counterpart in (IIc) (analyzed as [[Num [N A]] Dem]). However, the order in IIIc requires movement (as in [A N] [Dem [Num t_{A+N}]]) and it therefore does not have a mirrored counterpart (the structure *[[t_{A+N} Num] Dem] [N A] is generated through rightward movement of [N A], and so the order in (IVc) is unattested).

The Universal 20 pattern has been replicated in various other domains. Cinque (2009) discusses a range of phenomena that show that the same combination of fixed pre-head order and variable post-head order is found elsewhere when elements combined with a head must come in a strict hierarchy. Abels (2016) shows that the orders found in Germanic verb clusters match the orders found in the noun phrase (with the lowest element in the cluster corresponding to N), while Neeleman (2017) shows that the orders of PPs and the verb in Dutch do so, too. The ubiquity of the pattern suggests that the mechanisms that derive Universal 20 must be fundamental to the syntax.

The literature also contains papers that challenge aspects of Cinque's description of the typological data. These include Nchare (2012) and Dryer (2018).

Nchare shows that in Shupamem, a Grassfields Bantu tone language, nineteen orders are possible, well beyond the fourteen orders permitted by Cinque's system. Seven of these orders fall outside Cinque's typology, but strikingly, these seven orders are not neutral ones, but orders licensed by focus. It is of course a valid question how these marked orders are derived, but if Cinque's typology is limited to neutral orders, the Shupamem data are not necessarily problematic.

Dryer reports several exceptions to Cinque's hierarchy. Their import again depends on the criteria for inclusion into the dataset on which Cinque's typology is based. For example, Dryer reports that, alongside the

unproblematic Dem-Num-A-N order, Sierra Popoluca has Num-Dem-A-N

order, something highly problematic in Cinque's system and related proposals

if we are dealing with neutral order in a single extended projection. However,

Wendy Liz Arbey López Márquez (personal communication) reports that the

Num-Dem-A-N order conveys a partitive meaning, while the Dem-Num-A-N

order does *not* permit such an interpretation:

(4) a.   chiʔɨ   wɨsteen yɨʔp   joom   kuxtyat

        give.him two   this   new   sack

        'Give him two of these new sacks.'

    b.   yɨʔp   wɨsteen joom   kustyat ʔiniminy Peetu

        this   two   new   sack   brought   Peter

        'Peter brought these two new sacks.'

This suggests that the Num-Dem-A-N order may well involve two extended

nominal projections, on a par with English *two of these new sacks*. If the

numeral appears in the higher of these extended projections and the

demonstrative in the lower one, the order falls outside the data Cinque's system

is designed to capture.

    A related challenge concerns the issue of typological markedness. Any

survey of word order in the noun phrase, including Cinque's (2005, 2014) and

Dryer's (2018), has found that the typological distribution of orders follows a

power law, with some orders much more frequent than others. Such

asymmetries are captured by principles that determine markedness among

admissible orders. A difficult empirical question that then emerges is how we

can distinguish between highly marked orders and ungrammatical orders at the

tail end of the distribution, where data are scarce and some orders may hence be absent by chance or reported as real owing to misanalysis. The various approaches to Universal 20 draw the line in different places. Dryer (2018) does not rule out any order as ungrammatical, Steedman (2020) rules out two of the twenty-four logically possible orders, while Cinque (2005) rules out ten. Abels and Neeleman's (2012) proposal mirrors Cinque's for noun phrases with three modifiers but turns out to be more restrictive when structures with four modifiers are considered. This issue of where the boundary lies between highly marked and inadmissible orders remains unresolved and probably requires considerable additional research.

The main analytical disagreement within generative grammar involves the nature of the mirroring device. Cinque, who assumes Antisymmetry, argues that it takes the form of roll-up movement. Suppose that two phrases XP and YP (in the case at hand a pair taken from {Dem, Num, A}) can appear on either side of a lexical head L (in the case at hand N), showing a mirror image effect. Antisymmetry requires an underlying structure that is rightward descending. Specifiers of functional heads can host YP and XP in such a structure, making it possible to base-generate the YP-XP-L order. The reverse order is derived by moving LP across XP and then moving a category that dominates LP in its derived position across YP. To make this work and adhere to antisymmetric constraints on phrase structure, two additional functional projections must be postulated, whose specifiers function as landing sites for movement:

(5)   Insert Example 28-05 here

In (5), YP-XP-L is the base-generated order, YP-L-XP is derived if movement α but not movement β takes place, XP-L-YP is derived if β but not α takes place, and L-XP-YP results from a combination of both movements.

By contrast, Ackema and Neeleman (2002) and Abels and Neeleman (2009, 2012) assume that mirror image effects come about through variation in linearization. Consider the tree in (6), where linearization of nodes α and β is left open (α and β are nodes in the "spine" of the extended projection of the lexical head L). YP-XP-L order results if α follows YP and L follows XP, YP-L-XP is derived if α follows YP and L precedes XP, XP-L-YP is derived if α precedes YP and L follows XP, and L-XP-YP results from α and L preceding YP and XP, respectively.

(6)   Insert Example 28-06 here

There is a striking similarity between the two mirroring devices. Movements α and β in (5) and linearization of nodes α and β in (6) have the same function, showing that the two analyses have an identical parameter space. Moreover, Abels and Neeleman (2009) demonstrate that the two theories systematically generate representations with identical gross constituency, even when movement of (constituents containing) L is added to the mix. Thus, the trees generated by the two theories are isomorphic, in that XP, YP, L, and any traces of non-roll-up movements are grouped together in the same way. This means that it is difficult to find predictions that distinguish between the two approaches.

Nonetheless, Abels and Neeleman (2012) give three arguments against the antisymmetric approach, based on independently motivated constraints on movement: (i) antilocality, (ii) the ban on stranding of pied-piped material, and (iii) the A-over-A constraint. Each of these constraints can be shown to be incompatible with the extra movements required under Antisymmetry (that is, the movements not needed under an alternative that allows variation in the linearization of sister nodes). An additional piece of evidence against Antisymmetry is given in Neeleman (2017) and Neeleman and Payne (2020), where it is shown that the fact that antisymmetric trees are larger than their symmetric counterparts causes a paradox in the analysis of scalar focus particles (based on the observation that such particles must immediately c-command their semantic argument; see Barbiers 1995 and Neeleman and Van de Koot 2022a). These arguments are too intricate to review here, but in my reading of the subsequent literature they either stand unchallenged or have not been challenged successfully.

For this reason, I will henceforth assume that the language faculty allows variation in the linearization of sister nodes, acknowledging that many of the issues discussed have parallels within Antisymmetry.

28.3 Limitations on Mirroring

The Universal 20 paradigm requires that the position of each modifier with respect to the noun can be determined independently of the position of other modifiers. Nonetheless, this cannot be the whole story. If we look at the

statistical distribution of orders, we observe a clear, though imperfect, correlation between frequency and movement: Orders derived by movement are typically less frequent than orders that can be base-generated (exceptions include (IIId,f) in Table 28.2, which are more frequent than (IIb,d) in both Cinque's (2014) and Dryer's (2018) samples). If we put this factor to one side and limit ourselves to base-generated orders, it turns out that the two orders in which modifiers uniformly precede or uniformly follow the noun are more frequent than orders in which some modifiers are prenominal and others postnominal (this is again true in both Cinque's (2014) and Dryer's (2018) samples). In other words, there appear to be harmony effects within the extended nominal projection (where an extended projection is harmonic if the direction of headedness within it is uniform).

Harmony effects of this type are found elsewhere as well. For example, there is a strong correlation between the order of verb and object on the one hand and the placement of (low) adverbials and adpositional phrases on the other. The data in Table 28.3 are taken from Dryer (1991) (who also discusses other correlates of OV/VO order; see below). The measure used is the average of proportions of genera over areas. Thus, the 0.91 in the second cell of the second row indicates that in a random SOV language, there is a 91 percent chance that manner adverbs precede the verb.

Table 28.3

The correlation between OV/VO and modifier placement

	SOV	SVO
Manner adverb – V	0.91	0.25

Adpositional phrase – V   0.90   0.01

Thus, harmony within extended projections is not a grammatical principle, but

it is a powerful effect in terms of the statistical distribution of word orders.

An obvious way to think about harmony is in terms of the simplicity of

the grammars that generate the observed orders. For example, the N-A-Num-

Dem order can be generated by a grammar that has a single rule of linearization

that generalizes over all modifiers within the noun phrase. For concreteness'

sake, we may write this rule as in (7a). The order Num-N-A-Dem, however,

requires at least two rules. One can be a general rule identical to the one in (7a),

assuming the other rule specifically mentions NumPs, as in (7b). Competition

between the rules in (7b) is then regulated through the Elsewhere Condition,

which blocks application of the more general rule where the more specific rule

can apply.

(7)   a.   If N combines with XP and projects, then place N before XP.
      b.   If N combines with XP and projects, then place N before XP.
            If N combines with NumP and projects, then place NumP before N.

Hence, if simplicity of the target grammar is a factor in acquisition, harmony

effects may be modeled as a consequence of the way in which grammars are

learned. This view is backed up by experimental work by Jenny Culbertson and

collaborators (Culbertson et al. 2012; Culbertson and Newport 2015, 2017;

Culbertson et al. 2020). The finding of that work is that in artificial-language-

learning experiments, harmonic noun phrase word orders are preferred over

non-harmonic orders by both adults and children, irrespective of whether noun

phrases in the language they speak are themselves harmonic or non-harmonic.

However, harmony cannot be the sole driver of typological frequency. Consider the linearization systems in (8a,b), which are as complex as the ones in (7a,b) (again, the Elsewhere Condition regulates application of the rules in (8b)).

(8) a.   If N combines with XP and projects, then place XP before N.
     b.   If N combines with XP and projects, then place XP before N.
          If N combines with NumP and projects, then place N before NumP.

The systems in (8) would yield Dem-Num-A-N and Dem-A-N-Num order, respectively. If simplicity of grammar were the only factor that matters, we would expect these orders to be as frequent as the orders generated by (7a,b) (N-A-Num-Dem and Num-N-A-Dem). However, in both Cinque's and Dryer's samples, N-A-Num-Dem is clearly more frequent than Dem-Num-A-N, and Num-N-A-Dem is clearly more frequent than Dem-A-N-Num.

The point is strengthened by Culbertson et al.'s (2012) experimental work on Greenberg's (1963) Universal 18. Universal 18 states that "[w]hen the descriptive adjective precedes the noun, the demonstrative and the numeral, with overwhelmingly more than chance frequency, do likewise." This universal implies that A-N-Num and Num-N-A, although both non-harmonious, are not on a par. Indeed, the latter is more frequent than the former (by a factor of more than 4.5 in the World Atlas of Language Structures (WALS) sample that Culbertson et al. consider). Culbertson et al. show that in artificial-language-learning experiments learners prefer harmonic orders (Num-A-N and N-A-Num) over non-harmonic orders (Num-N-A and A-N-Num). However, they

also disfavor A-N-Num in comparison with Num-N-A. The distinction is characterized as discrete and decisive.

There is not much insight into what factors other than harmony may influence the typological distribution of different word orders. For example, Dryer (2018) suggests that there may be a preference of N-A order over A-N order. However, he has no account of *why* this preference should be there. Culbertson et al. (2012) suggest a possible theoretical explanation for Universal 18 based of the Final-Over-Final Constraint (see below), but thate explanation is unlikely to be correct, as it ~~incorrectly~~ mistakenly treats adjectives as heads in the extended nominal projection (Svenonius 1994; Matushansky 2002; Cinque 2010) and does not extend to the contrast between Dem-Num-A-N and N-A-Num-Dem. Another possible explanation discussed by Culbertson et al. relies on parsing efficiency. This explanation is more likely to be correct but requires further investigation. Thus, in this domain we have data, but ~~no convincing account~~limited understanding.[1]

In addition to harmony effects *within* extended projections, there are also harmony effects *between* extended projections. For example, there is a strong positive correlation between the placement of the verb with respect to its complement(s) and the placement of adpositions with respect to theirs. Table 28.4 gives information from Dryer (1991). As before, the measure used is the average of proportions of genera over areas. Thus, the chance that a random

---

[1] For more complete discussions of the factors that determine the typological frequency of U20 orders, see Dryer 2018 and Abels and Neeleman 2023.

SOV language is postpositional is 96 percent, while the chance that a random

SVO language is comes out much lower, at 14 percent.

Table 28.4

The relation between OV/VO and adpositional placement

	SOV	SVO
Postpositions	0.96	0.14

As in the case of harmony effects within extended projections, the

correlation in Table 28.4 may be explained in terms of the simplicity of the

grammars involved. This requires linearization rules that can generalize over

categories. For example, an SOV language with postpositions can be described

as having the linearization rule in (9a), with X ranging over P and V. By

contrast, a prepositional SOV language would have to have at least the rules in

(9b) (with the Elsewhere Condition blocking the first rule where the second can

apply).

(9) a. If X selects a DP and X projects, then place DP before X.
     b. If X selects a DP and X projects, then place DP before X.
         If P selects a DP and P projects, then place P before DP.

Dryer (1991) discusses a number or other correlates of OV/VO order. These

include the order of a predicate and a copula, the order of subordinators like

'because' with respect to the clause they introduce, the order of non-affixal

plural heads with respect to the noun, and the order of auxiliaries with respect

to V.

Given the existence of cross-categorial harmony, one would expect the

structures in (10a) and (10b) to be more frequent than those in (10c) and (10d)

(for instance when X is an auxiliary and Y a verb).

As a statistical statement, this is almost certainly correct. However, there is a body of work suggesting that a statistical statement is not enough. It has been claimed that there are specific circumstances in which disharmonious orders are not just rare but ungrammatical.

Perhaps the best-known proposal in this domain is the Final-Over-Final Constraint (FOFC) (Biberauer et al. 2007, 2008, 2014; Sheehan et al. 2017), which states that "final over initial orders" are ruled out. The FOFC expresses a contrast between (10c) and (10d). While (10c) is grammatical in principle (though dispreferred), (10d) is ruled out altogether if X and Y are heads in the same extended projection. The proposal in Biberauer et al. (2014) is formulated in antisymmetric terms but can easily be rephrased in terms of a more traditional view of linearization. To begin with, assume that functional heads by default precede their complements, and that for the complement to precede, the functional head must bear a special feature triggering non-default order. In (11), this feature is marked by the diacritic > (intended to be read as an arrow pointing to the right).

(11)   a.   If X selects YP, X projects, and X and YP belong to the same extended projection, then place X before YP.

         b.   If $X^>$ selects YP, $X^>$ projects, and $X^>$ and YP belong to the same extended projection, then place YP before $X^>$.

If so, the FOFC can be formulated as in (12).

(12)   **Final-Over-Final Constraint**
       If a head X_i in the extended projection of a lexical head L carries >, then so does X_{i+1}, where X_{i+1} is c-selected by X_i in L's extended projection.

Assuming that X and YP in (10) belong to the same extended projection, (12) rules out (10d) as follows. (i) As X follows YP in (10d), it must carry >. (ii) If X carries >, then Y must carry > as well. (iii) Hence, Y should follow ZP. (iv) However, Y precedes ZP.

The alternative formulation of the FOFC in (12) closely follows the original in Biberauer et al. (2014) and seems identical in its predictions regarding word order. Hence, the FOFC, if correct, does not provide evidence for Antisymmetry. That, of course, does not make it any less interesting as a potential restriction on the linearization of syntactic structures.

The evaluation of the FOFC as a potential universal mainly hinges on what limits should be assumed regarding the domain in which it holds, an issue only partly resolved. There are data confirming the asymmetry between (10c) and (10d). In Germanic, for example, the orders in (13a–c) are attested, but the FOFC-violating order in (13d) is not (Biberauer et al. 2014 and references mentioned there). A similar pattern is found in languages that have both complementizers and particles expressing polarity. Sheehan (2017) notes that in the World Atlas of Language Structures, the order in (14d) is very much rarer than the orders in (14a–c). The order in (14a) is found in 72, (14b) in 45, and (14c) in 74 languages. The order in (14d) occurs in only four languages (for related discussion, see Dryer 2012: 76 and Cinque 2023, §5.3).

(13) a. [Aux [V O]]
      b. [[O V] Aux]
      c. [Aux [O V]]
      d. *[[V O] Aux]

(14) a.   [C [Pol VP]]
      b.   [[VP Pol] C]
      c.   [C [VP Pol]]
      d.   *[[Pol VP] C]

The definition in (12) restricts application of the FOFC to heads in the same extended projection. However, it is not obvious that this restriction makes the right cut.

On the one hand, there are structures that display the FOFC effect, but that arguably do not constitute single extended projections. A case in point is affixation. Hein and Murphy (2022) argue that suffixes must attach to a head-final base, so that (15b) is ruled in, but (15d) is excluded (the same claim can be found in Ackema and Neeleman 2004, which does, however, identify several types of counterexamples). The effect is so strong that bases that normally have a head-initial order sometimes assume a head-final order under suffixation (cf. *stand out – outstanding*).

(15) a.   [Aff$_X$ [$_Y$ Y Z]]
      b.   [[$_Y$ Z Y] Aff$_X$]
      c.   [Aff$_X$ [$_Y$ Z Y]]
      d.   *[[$_Y$ Y Z] Aff$_X$]

If the ungrammaticality of (15d) has its source in the FOFC, we expect an asymmetry with (15c), which is equally disharmonic, but does not violate the FOFC. Hein and Murphy argue that indeed there is such an asymmetry, basing themselves on a typological survey of languages with overt nominalizers. The statistics are as follows: [NMLZ [V O]] is attested in five languages, [[O V] NMLZ] in ten languages, and [NMLZ [O V]] in one language. There are no attestations of *[[V O] NMLZ]. (There are also three languages with O-NMLZ-V

order in the sample; Hein and Murphy analyze those as derived from [NMLZ [O V]].)

The argument is somewhat weakened by the limited number of languages showing NMLZ-O-V order. Be this as it may, if Hein and Murphy are correct in their assessment of the data, the scope of the FOFC, as defined in (12), is too narrow and must be extended to at least nominalizations. They suggest this should be possible on the assumption that the FOFC applies to extended projections if nominalizations are "mixed categories" that are simultaneously nominal and verbal. I am skeptical of this assessment, as there are good empirical and theoretical reasons to reject this analysis of nominalization (Ackema and Neeleman 2004, §5.9 and references mentioned there).

On the other hand, there are particles in the verbal extended projection that appear to violate the FOFC. Zeijlstra (~~2022~~2023) mentions VO languages with clause-final tense, mood and aspect particles, negative particles, and interrogative particles. Not all of these constitute straightforward counterexamples. Some may be phrases rather than heads. However, there is certainly a challenge here, one that suggests that more work on the FOFC's domain of application is necessary. As an alternative, Zeijlstra~~'s~~ (~~2022~~2023) proposes that the FOFC only applies to heads connected by movement in at least some instantiations of the pertinent structure (see section 28.2 and section 28.4). His derivation of this fact starts from the hypothesis that non-string-vacuous head movement must be leftward, so that a head-initial phrase cannot

be dominated by a head-final one if the lower head moves to the higher head. A head-final phrase can always be dominated by another head-final phrase, as any movement between the two heads would be string-vacuous. Thus, the FOFC receives a different derivation and hence a different domain of application.

The FOFC is not the only constraint that rules out or suppresses disharmonic orders. Philip (2012, 2013) suggests that there is a certain class of heads, known as linkers, that are subject to a stronger constraint, which she dubs the Head Proximate Filter (see also Frazier's (1980) Head Adjacency Constraint, Rijkhoff's (1984) Principle of Head Proximity, and Hawkins' (1990, 1994) Principle of Early Immediate Constituents). Linkers are elements that mark an independently existing relationship between a head and one of its dependents (Dik 1983; Zwart 2009; Philip 2012; Den Dikken 2006). They are systematically attached to the dependent category and are typically realized between the dependent and the head, as schematized in (16) for the case where the head is a noun (Dik 1983, 1997).

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

Philip formulates the Head Proximate Filter as a violable constraint. However, her account is organized in such a way that linkers must systematically satisfy it (unless the linker–dependent combination undergoes movement). This predicts the distribution in (16) for linkers in the nominal domain.

The highest head in the extended projection of a dependent must be linearly contiguous with the head it is in relationship with.

Philip uses a typologically balanced sample of eighty-three languages to motivate this distribution. The numbers are as given in (18).

(18) a.   [[XP LNK] N]:   32 languages
      b.   [N [LNK XP]]:   50 languages
      c.   *[[LNK XP] N]:   potentially 1 language
      d.   *[N [XP LNK]]:   0 languages

Within the verbal domain, complementizers can arguably be classified as linkers, and they indeed show a distribution parallel to their nominal counterparts (data from Dryer 2009; see also Hawkins 1994, Bayer 1996 and Cinque 2005).

(19) a.   [[TP C] V]:   12 languages
      b.   [V [C TP]]:   157 languages
      c.   *[[C TP] V]:   0 languages
      d.   *[V [TP C]]:   0 languages

The Head Proximate Filter differs from the FOFC in two ways. First, it does not discriminate between disharmonic orders: Both "initial over final" and "final over initial" orders are ruled out. In contrast, the FOFC rules out "final over initial" orders but tolerates "initial over final" orders. Second, at least in the domain of linkers, the Head Proximate Filter regulates order across extended projection boundaries, while FOFC effects are limited to extended projections, at least according to Biberauer et al. (2014).

28.4 Dependencies

The discussion so far mainly concerned "local" linearization, involving structural sisters or at least heads local to each other in the structure. However, there are also linearization constraints on the "antecedent" and "dependent" in dependencies that span larger distances. For example, if we put reconstruction effects to one side, anaphors and bound pronouns (which are dependents) tend to follow the category that binds them (their antecedent). In the same vein, movement is typically (though not exclusively) leftward, so that the moved category (the antecedent) precedes the trace (the dependent). Observations of this type raise two questions. (i) Is the effect absolute or merely a tendency? (ii) Do the data require a grammatical explanation or does the effect have an extra-grammatical source? Where a grammatical explanation is necessary, there are additional issues of implementation.

The left–right asymmetry found with grammatical dependences is perhaps most convincing in the case of head movement. As mentioned in section 28.2, the Universal 20 data in Cinque (2005) and subsequent work require that there is no rightward movement of (constituents containing) the noun in the extended nominal projection. The fact that the Universal 20 pattern has been replicated in other domains suggests that there, too, rightward head movement (including movement pied-piping other material) should be ruled out.

There are two possible explanations of this fact. Antisymmetry implies that syntactic representations are uniformly rightward descending, so that a

head attached higher in an extended projection precedes heads lower in that extended projection. Consequently, head movement must be leftward. An alternative extra-grammatical explanation could be based on processing. If parsing involves immediate structure assignment (Gorrell 1995), and movement is dealt with using a filler-driven strategy (Phillips and Wagers 2007), it follows that there is a fundamental difference between leftward and rightward movement: The former involves insertion of a trace while the structure is being built, while the latter involves insertion of a trace in an already built structure. Consequently, leftward movement will be easier to parse than rightward movement. It stands to reason that grammars that are harder to parse are harder to learn, so that, over time, grammars with rightward movement are at a disadvantage compared to grammars with leftward movement (or no movement at all). Even if this pressure is small, it could result, given enough time, in the disappearance of languages with rightward movement.

This is not enough, however, as it would rule out standard rightward phrasal movement in addition to rightward movement of (constituents containing) the head of an extended projection. As discussed below, there is robust evidence that the former can be rightward. Hence, any account based on processing must distinguish between these two types of movement. One aAttempt to explain why movement affecting the head of an extended projection must be leftward can be found in Ackema and Neeleman (2002) and Abels and Neeleman 2023. An interesting promising alternative is the account

by Medeiros (2018), who uses a version of Knuth's (1968) stack-sorting algorithm to map surface word orders to an underlying base structure. As Medeiros shows, this algorithm captures Universal 20 effects quite straightforwardly. Hence, incorporating it into the parser as a mechanism for analyzing neutral word orders (including orders derived by head movement) may remove the need for a grammatical ban on rightward head movement, although the algorithm could also be argued to motivate such a grammatical constraint. I should note that Medeiros goes much further than this, suggesting a model in which the parser *is* the grammar, a hypothesis that I cannot evaluate in this chapter.

There are clear asymmetries between regular leftward and rightward phrasal movement, with the latter typically more limited and more local than the former. Nonetheless, a convincing case can be made for the existence of rightward phrasal movement. Overfelt (2015) shows, following Larson (1989) and Nissenbaum (2000), that (in English) rightward displacement of a DP beyond standard heavy NP shift must result in the binding of a parasitic gap by the displaced DP. As long as each modifier crossed by the movement contains a parasitic gap, the rightward shift of the DP is in principle unbounded, suggesting that the movement is triggered by the need to license a parasitic gap. What this study shows, then, is that while rightward and leftward movement may span similar distances, the *trigger* for rightward phrasal movement may be quite different from the triggers assumed for leftward phrasal movement.

As second asymmetry in phrasal movement has to do with CPs. As is well known, CPs that appear to move leftward in English must always be related to a position in which NPs are allowed, something that is not true of CPs that move rightward. For example, *attribute* can take a DP-complement, but *reason* cannot. Hence, *attribute* permits CP fronting, but *reason* does not (data from Alrenga 2005).

(20) a. *That the Giants would probably win the World Series, (I think that) most baseball fans reasoned.
      b. That these consonants behave exceptionally, we can attribute to the fact that they are coronals.

One account for this observation is that leftward-displaced CPs are base-generated in their surface position, and that what moves is a silent nominal operator (Koster 1978; Alrenga 2005; Moulton 2013). Another possibility is that fronted CPs are dominated by a silent DP layer (Takahashi 2010). Either way, the picture that emerges is one in which CPs can move rightward, but not leftward, as argued explicitly in Bruening (2018) and Bruening and Al Khalaf (2020). This contrasts with the behavior of DPs, which can move in either direction (as just discussed, there are examples of unbounded rightward DP movement).

It is not clear why CPs should behave in this way. One possible approach is to capitalize on the facts outlined in (19), namely that complementizers, like other linkers, have a strong tendency to be realized between the clause that they introduce and the head of the structure in which that clause is embedded. Rightward movement of CP preserves that order, but

leftward movement does not. If this is on the right track, one would expect that

CPs in languages with head-final complementizers *can* be moved leftward.

Linear constraints have also been proposed for relations of anaphoric

dependence. One such constraint is Williams' (1997) General Pattern of

Anaphoric Dependence (GPAD; see also Langacker 1969 and Ross 1969).

Williams argues that in an anaphoric dependency the dependent category must

either follow its antecedent or be contained in a clause subordinate to that

antecedent. The following data, taken from Williams' paper, illustrate this. (In

(21d), *term paper* is stressed to avoid accommodation. If destressed, it could be

anaphoric on an earlier mention of *term paper*.)

(21) a.   Anyone [who has written his term paper] can turn it in to me now.
      b.   Anyone [who has written it] can turn his term paper in to me now.
      c.   Anyone can turn his term paper in to me now [who has written it].
      d.   *Anyone can turn it in to me now [who has written his term paper].

The GPAD applies quite broadly. For example, VP ellipsis appears to be

subject to it:

(22) a.   People [who claim that they understand quantum mechanics] never do e_{VP}.
      b.   People [who claim that they do e_{VP}] never understand quantum mechanics.
      c.   People never understand quantum mechanics [who claim that they do e_{VP}].
      d.   *People never do e_{VP} [who claim that they understand quantum mechanics].

The GPAD implies that in the absence of subordination of the dependent

category, the latter must follow its antecedent. In (21a,d), it is the antecedent

that is subordinated and hence this category must precede the pronoun *it*, a

requirement met in (21a), but not in (21d). Similarly, when the antecedent and

dependent are contained in different sentences (so that neither is subordinate to

the other), coreference is predicted to require precedence of the antecedent.

This is correct for examples like (23), though see Carden (1982) on "withheld antecedent discourse structure").

(23)   a.   John walked in. He wore a hat.
       b.   *He walked in. John wore a hat.

Crucially, the GPAD requires precedence of the antecedent when it is a clausemate of the dependent category. This seems correct, but with some qualifications. While there is a sharp contrast between (24a) and (24b), fronting of the constituent containing the pronoun yields a structure in which backward anaphora is possible, suggesting that if linear order is at play, it is the underlying linear order.

(24)   a.   John's mother loves his father.
       b.   *His mother loves John's father.
       c.   It's his father that John's mother loves t.

Williams argues that, like coreference, variable binding is subject to the GPAD, offering data like (25) in evidence. Notice that, as expected, a bound-variable pronoun can precede the quantifier if embedded in a subordinate clause:

(25)   a.   *His girlfriend loves every British soldier.
       b.   [That he might someday meet the queen] inspires every British soldier.
       c.   [That an enemy sniper shot him] bothered every soldier in the hospital.

The pattern extends to dative constructions (Bruening 2001; Janke and Neeleman 2012).

(26)   a.   *I returned his manuscript to every author.
       b.   I gave a flower [that Peter said she would like] to every girl in my class.
       c.   I explained [how an enemy sniper had shot him] to every soldier in the hospital].

An application of the GPAD to variable binding fits well with the recent resurgence of precede-and-command as an alternative to c-command for some dependencies. Barker (2012), for example, argues that surface c-command is

too restrictive a notion to account for the full range of variable binding data in English and should be replaced by a looser relation (scope plus (reconstructed) linear order). There is a range of further work in the same spirit (Jackendoff 1990; Ernst 1994; Bruening 2014; Neeleman and Van de Koot 2022b; for experimental evidence, see Moulton and Han 2018; for discussion of linear order effects in coordination, see Bruening and Al Khalaf 2016 and Nevins and Weisser 2018). While the claim that linear order plays a role in binding is shared by these authors, there are disagreements regarding the structural notion of command, something that I will not explore here.

As with movement, there is a question as to what explains linear effects in binding, and as with movement, there is a body of work suggesting that processing must at least be part of the answer. Shan and Barker (2006), Barker and Shan (2008), and Barker (2009, 2012) attribute the precedence constraint on variable binding to a requirement to evaluate a quantifier (in a specific technical sense) *before* any pronoun that it binds. This explanation of left–right effects in terms of time is clearly indicative of processing. Bruening (2014) adopts a similar approach, arguing for a left-to-right processing model sensitive to phasal nodes.

More can be said about linear effects in dependencies than fits in this chapter. However, to return to the questions raised at the start of this section, there seems to be evidence that linear constraints hold of at least some dependencies. It is easy to see that these constraints contribute to parsing efficiency, as they favor linear orders in which dependent categories follow the

antecedents that they are dependent on. What is much less clear is whether the relevant constraints are instances of grammatical requirements that reduce pressure on the parsing process, or whether they result from the gradual pruning in language change of grammars that are more difficult to parse and hence more difficult to learn. The answer to this question may not be the same for the various linear constraints the literature provides evidence for. Finally, to the extent that the constraints under discussion are grammatical, they raise thorny questions of implementation. For example, suppose that the constraint on head movement is grammatical. Does this mean that the head movement operation itself is subject to a linear requirement or is it possible to design a linearization rule that places the moved category to the left of the category containing the trace?. A rule of the latter type is easier to imagine if the category containing the trace carries some sort of movement diacritic akin to the slash features of HPSG.

28.5 Compactness

A final phenomenon that may necessitate reference to linear order consists of adjacency requirements. Perhaps the best-known generalization relevant here is that SOV languages like Dutch allow adverbs to surface between the verb and its internal arguments, while such adverbial intervention is prohibited in English, as illustrated in (27). The ban on adverbial intervention in SVO languages has been dubbed *compactness* by Haider (2005, 2014).

(27) 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 pattern in (27) extends beyond Dutch and English. The Adv-O-V/O-Adv-

V alternation can be observed in Afrikaans, Armenian, Assamese, Basque,

Bengali, Frisian, German, Georgian, Hindi, Japanese, Kannada, Kiowa,

Korean, Lezgian, Malayalam, Pashto, Persian, Quechua, Sakha, Sami, Tatar,

Tsez, Turkish, Uyghur, and Uzbek. Indeed, I do not know of any ~~no~~ SOV

languages with preverbal adverbials that ban adverbial intervention. By

contrast, at least some SVO languages with postverbal adverbs are like English

in permitting V-O-Adv, but not V-Adv-O order (except through heavy-XP

shift~~.~~). Bari and Tuki are examples, as shown in (28); similar data can be found

in Edo, Igbo, Khmer, Maybrat, Thai, Vietnamese, and Xhosa.

(28) a. Teleme   a   kop   ⟨*'de'de⟩   kene   ⟨'de'de⟩.   Bari (Creider 1989)
          monkey   TNS   catch   quickly   branch   quickly
          'The monkey caught the branch quickly.'
      b. Vakútu   vá-mu-será   ⟨*ísími⟩   mbasá   ⟨ísími⟩.   Tuki (Biloa 2013)
          women   SM-PL-sell   rapidly   corn   rapidly
          'The women sell the corn rapidly.'

There are SVO languages in which adverbs may surface between the verb and

its internal arguments because the verb moves. Icelandic is a case in point (see

(29a)). Such languages can nonetheless provide evidence for compactness. In

Icelandic, verb movement is suppressed in the presence of an auxiliary, and as

it turns out, compactness can be observed with in-situ verbs (see (29b)); similar

data can be found in Danish, Norwegian, and Swedish.

(29) a. Hann   missti   ⟨skyndilega⟩ t_V takið   ⟨skyndilega⟩.   Icelandic (Jóhanne
          he.NOM   missed   suddenly   the.grip.ACC   suddenly

‘He suddenly lost his grip.’

    b.   Hann   hafði ⟨skyndilega⟩ t_{AUX} misst ⟨*skyndilega⟩ takið ⟨skyndi

        he.NOM had suddenly   lost suddenly   the.grip.ACC suddenl

        ‘He had suddenly lost his grip.’

A related pattern is found in Aghem, which has a focus position immediately

after the verb (see Hyman 2010). I assume, for concreteness' sake, that focused

elements move to a position at the left edge of the VP and that the verb

subsequently moves as well, giving rise to the observed verb-focus order.

Crucially, adverbs can only precede the object when narrowly focused, as in

(30a). In the absence of such narrow focus, Aghem provides evidence for

compactness, as shown in (30b).

(30)   a.   tí-bvú   tì-bìghà   mô   zì   **nɛ́**   [t_V kí-bɛ́ t_{Adv}].   Aghem

            dogs   two   P1   eat   today   fufu

            ‘the two dogs ate fufu **today**’

    b.   tí-bvú   tì-bìghà   mô   zì   ⟨*nɛ́⟩ kí-bɛ́ ⟨nɛ́⟩.

            dogs   two   P1   eat today fufu   today

            ‘The two dogs ate fufu today’

Similar data can be found in Bemba (Costa and Kula 2008), Makhuwa (Van

der Wal 2006), and Zulu (Cheng and Downing 2012).

    Sometimes, the evidence for compactness is quite indirect. Suppose that

the verb moves, but there are no adverbial positions between its trace and the

object (as compactness implies). In that case, if two adverbs intervene between

the verb and the object, the higher adverb must precede the lower adverb (see

(31a)). By contrast, if there were adverbial positions between V and O (in a

non-compact verb-initial VP), an intervening higher adverb should be able to

follow an intervening lower one, whether the verb moves or not (see (31b)).

(31) a.   [S [V [Adv₂ [Adv₁ [tᵥ O]]]]]
      b.   *[S [[[V Adv₁] Adv₂] O]]

The following data (32) thus confirm that Czech, even though it permits adverbial intervention, has a compact VP (note that 'again' resists being in the scope of 'three times' and that time adverbs tend to attach higher than manner adverbs).

(32) a.   Petr   si   přečetl   ⟨znovu⟩   třikrát   ⟨*znovu⟩ dopis.   *Czech* (Jir
           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*)

The pattern found in Czech extends to Italian (Cinque 1999), Polish (Natalia Cichosz, p.c.), French (Abeillé and Godard 2003), and Slovenian (Nadja Rajgelj, p.c.). I am not aware of SVO languages in which there is convincing evidence for a non-compact VP, although there are many such languages whose status is unclear. (Within the wider class of languages with a head-initial VP (VSO and VOS), there is one that arguably has a non-compact VP, namely Malagasy; see Pearson 2007.).

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. This notion of *impenetrability* is already present in Chomsky (1965), where a monolithic VP is dominated by a category labeled PredP in which adverbs are attached. The same idea is inherent in Chomsky's (1986) sisterhood condition on θ-role assignment. If an adverb

were combined with the verb before the object was attached, the verb and the object would not be sisters and hence the object could not be θ-marked. Thus, the verb–object combination is impenetrable.

The second approach assumes that the verb must be adjacent to certain DP arguments. This idea goes back to Keyser (1968), but its best-known implementation is due to Stowell (1981), who argues that the verb must be adjacent to any DP argument it case-marks, a constraint known as *case adjacency*.

Both impenetrability and case adjacency explain why adverbs cannot separate a verb from its object(s) in English. However, they shed no light on the grammaticality of adverbial intervention in SOV languages like Dutch. Such languages therefore require the introduction of a movement operation, often referred to as *scrambling*, which shifts the object leftward, away from the verb (Kerstens 1975; Hoekstra 1984). The proposal comes with the implicit assumption that English (and other SVO languages) lack a parallel rightward movement operation.

Neither impenetrability nor case adjacency can cope with the distribution of very low adverbs in English. As the Dutch example in (33a) shows, there are certain adverbs that must appear in a position lower than the object. As (33b) shows, such adverbs must follow the object in English; they cannot appear preverbally (an observation reported as early as Jackendoff 1972).

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

   b. Ava ⟨*badly⟩ treated Carlos ⟨badly⟩.

In other words, there is incursion into the English VP by adverbs that follow the object, but not by adverbs that precede the verb. Impenetrability does not allow any adverb incursion into the VP, and therefore cannot explain contrasts like the one in (33b). Case adjacency only cares about adverbs that separate the object and the verb. It hence allows low attachment of adverbs whether they precede the verb or follow the object, with the consequence that the contrast in (33b) again remains unexplained.

There is consensus that low adverbs should be analyzed as occupying the lowest position in a VP-shell structure (Larson 1988; Vanden Wyngaerd 1989, among others). This analysis has the benefit that it explains why preverbal adverbs are attached relatively high. If the verb moves out of VP, any preverbal adverb must be VP-external as well. Thus, if we assume that the verb always moves, contrasts like the one in (33b) fall into place. However, verb movement routinely crosses adverbials, so VP-shell accounts make it harder to capture the fact that no adverbs may intervene between the verb and its objects in SVO languages.

There are several attempts to explain the pattern in grammatical terms (Chomsky 1995; Saito and Fukui 1998; Neeleman and Weerman 1999; Haider 2005, 2014; Janke and Neeleman 2012, among others). I cannot evaluate these here, but in my assessment none of them covers the full set of relevant data.

However this may be, processing may play a role in this domain, too (cf. Hawkins 2008). Suppose that the processing of verb–argument relationships is argument-driven, much like the processing of movement dependencies is filler-driven (this is likely as the argument structure of verbs is highly variable). Then, if an argument precedes the verb, the theta-grid of the verb can be built to accommodate the argument as the structure-building process continues. If the argument follows the verb, however, the parser needs to re-access its left context to adjust the verb's theta-grid, so that the argument can be accommodated. This will necessarily lead to a pause in the structure-building process. It is therefore advantageous to minimize the distance between verb and object in VO languages, but object–verb adjacency has no benefits in OV languages beyond a general minimization of the length of the dependency. Thus, the grammatical constraints that capture the contrast between SOV and SVO languages regarding compactness (whatever they are) may well find their motivation in processing.

I should point out two further aspects of the problem. First, adjacency phenomena of this type are not limited to the verb and its objects. It has been noted by Giurgea (2009), Adger (2012), and Belk and Neeleman (2017) that attributive APs precede certain other categories (PPs, genitive DPs, and so on), when the noun precedes both (see also Cinque 2010, §6.2). However, in languages in which PPs and APs precede the noun, both PP-AP-N and AP-PP-N orders are attested. Thus, there appears to be an N-AP adjacency effect that is like the verb-object adjacency effect discussed above. However, accounting

for the two effects in the same terms is challenging, given that the grammatical functions of objects and attributive adjectives are not comparable.

Second, there is a certain unresolved tension between Universal 20, which insists on fixed order preceding a given head and variable order following it, and the issue of compactness/adverbial intervention, which shows variable order preceding the verb and fixed order following it. Of course, Universal 20 is a typological generalization about word-order variation between languages, while the issue of compactness/adverbial intervention has to do with word-order variation within a single language. However, this is not enough to side-step the issue, as there are examples of Universal 20-style variation that involve language-internal alternations. One of these concerns so-called PP-over-V in Dutch (Koster 1974; Barbiers 1995; Neeleman 2017). Another has to do with the distribution of particles. As is well known, these systematically appear adjacent to the verb in the Germanic SOV languages (modulo verb second), while in the Germanic SVO languages they show variable behavior. In Danish, particles must appear between object and verb, in Swedish, they must follow the object, while in English and Norwegian, both options are available. This pattern is reminiscent of the Universal 20 typology (fixed order preceding the head and variable order following it). If this is not accidental, the variation in particle placement in English and Norwegian again suggests that the Universal 20 pattern has language-internal instantiations, an interesting conclusion given that the data are found in the same domain that its

mirror image, the compactness/adverbial intervention generalization, pertains to.

28.6 Concluding Remarks

As mentioned at the outset of this chapter, one of the lasting achievements of generative grammar is its demonstration of the structure-dependence of grammatical rules. But an acknowledgment of structure dependence does not imply a refutation of linear conditioning. The two issues are logically independent. I have shown in this chapter that there are various syntactic phenomena that seem to require reference to linear order, ranging from the harmony of projection to the leftward orientation of movement and variable binding and from the FOFC and head proximity effects with linkers to verb–object adjacency in SVO languages. It is an unresolved question what to do with observations of this type. In particular, it is not clear whether the syntax should be allowed to refer to linear order or whether the effects can be captured in other ways.

At least a subset of harmony effects can perhaps be captured as by-products of the acquisition process (see Chapter 9, section 9.3, for exposition on the acquisition of word order). The linearization system in harmonic grammars is likely to be simpler than that in disharmonic grammars, and this may give rise to statistical effects given enough time. However, other observations related to word order harmony cannot be captured using just simplicity of the grammars involved. For example, the asymmetry inherent in

the FOFC and the requirement of intervention in the placement of linkers appear to require a grammatical statement of some sort, at least in first approximation. The statements in question could be syntactic or mapping rules associating the syntax to linear order at the PF interface, with currently little evidence pointing one way or the other.

Linear constraints on dependencies such as head movement and variable binding, as well as the fact that many if not all SVO languages have a compact VP, may also have an extra-grammatical motivation, namely in parsing. This raises the question whether the parsing explanation is enough, meaning that no grammatical constraints need to be invoked in accounting for the data, or whether parsing considerations are merely the motivation behind grammatical constraints that rule out structures with rightward head movement or V-Adv-O order. Certainly, the effects are stronger than found with, say, garden path phenomena, which means that either the grammar must be invoked, or the parser must be organized to yield more absolute effects.

Abbreviations (outside of the Leipzig Glossing Rules)

P$_1$	today past	SPF	specific
SM	subject marker	TNS	tense

chapter-references

References

Abeillé, A., and D. Godard. 2003. French adverbs without functional projections. In M. Coene et al. (eds.), *Current Studies in Comparative Romance Linguistics*, 1–39. Amsterdam: John Benjamins.

Abels, K. 2016. The fundamental left–right asymmetry in the Germanic verb cluster. *Journal of Comparative Germanic Linguistics* 19: 179–220.

Abels, A., and A. Neeleman. 2009. Universal 20 without the LCA. In J. M. Brucart et al. (eds.), *Merging Features: Computation, Interpretation, and Acquisition*, 60–79. Oxford: Oxford University Press.

Abels, K., and A. Neeleman. 2012. Linear asymmetries and the LCA. *Syntax* 15: 25–74.

Abels, K., and A Neeleman. 2023. e. Ms. UCL (https://ling.auf.net/lingbuzz/007739).

Aboh, E. 2004. *The Morphosyntax of Complement-Head Sequences*. Oxford: Oxford University Press.

Ackema, P., and A. Neeleman. 2002. Effects of short-term storage in processing rightward movement. In S. Nooteboom et al. (eds.), *Storage and Computation in the Language Faculty*, 219–256. Dordrecht: Kluwer.

Ackema, P., and A. Neeleman. 2004. *Beyond Morphology: Interface Conditions on Word Formation*. Oxford: Oxford University Press.

Adger, D. 2012. *A Syntax of Substance*. Cambridge, MA: MIT Press.

Alrenga, P. 2005. A sentential subject asymmetry in English and its implications for complement selection. *Syntax* 8: 175–207.

Barbiers, S. 1995. *The Syntax of Interpretation*. PhD dissertation, University of Leiden.

Barker, C. 2009. Reconstruction as delayed evaluation. In E. Hinrichs and J. Nerbonne (eds.), *Theory and Evidence in Semantics*, 1–28. Stanford, CA: CSLI Publications.

Barker, C. 2012. Quantificational binding does not require c-command. *Linguistic Inquiry* 43: 614–633.

Barker, C., and C.-C. Shan. 2008. Donkey anaphora as in-scope binding. *Semantics and Pragmatics* 1: 1–41.

Bayer, J. 1996. *Directionality and Logical Form*. Dordrecht: Kluwer.

Belk, Z., and A. Neeleman. 2017. AP-adjacency as a precedence constraint. *Linguistic Inquiry* 48: 1–45.

Berwick, R. C., and N. Chomsky. 2016̶7̶. *Why Only Us?* Cambridge, MA: MIT Press.

Biberauer, T., A. Holmberg, and I. Roberts. 2007. Disharmonic word-order systems and the Final-Over-Final Constraint (FOFC). In A. Bisetto and F. Barbieri (eds.), *Proceedings of Incontro di Grammatica Generativa 33*, 86–105.

Biberauer, T., A. Holmberg, and I. Roberts. 2008. Structure and linearization in disharmonic word orders. In C. B. Chang and H. J. Haynie (eds.), *Proceedings of the Western Coast Conference on Formal Linguistics 26*, 96–104.

Biberauer, T., A. Holmberg, and I. Roberts. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45: 169–225.

Biloa, E. 2013. *The Syntax of Tuki*. Amsterdam: John Benjamins.

Bruening, B. 2001. QR obeys superiority: ACD and frozen scope. *Linguistic Inquiry* 32: 233–273.

Bruening, B. 2014. Precedence-and-command revisited. *Language* 90: 342–388.

Bruening, B., and E. Al Khalaf. 2016. Linear effects in ATB movement. In P. Farrell (ed.), *Proceedings of the Linguistics Society of America 1*, article 10.

Bruening, B., and E. Al Khalaf. 2020. Category mismatches in coordination revisited. *Linguistic Inquiry* 51: 1–36.

Carden, G. 1982. Backwards anaphora in discourse context. *Journal of Linguistics* 18: 361–387.

Cheng, L., and L. Downing. 2012. Against FocusP: Evidence from Durban Zulu. In I. Kučerová and A. Neeleman (eds.), *Contrasts and Positions in Information Structure*, 247–266. Cambridge: Cambridge University Press.

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press

Cinque, G. 1996. The "Antisymmetric" Programme: Theoretical and typological implications. *Journal of Linguistics* 32: 447–464.

Cinque, G. 1999. *Adverbs and Functional Heads*. Oxford: Oxford University Press.

Cinque, G. 2005. Deriving Greenberg's Universal 20 and its exceptions. *Linguistic Inquiry* 36: 315–332.

Cinque, G. 2009. The fundamental left-right asymmetry of natural languages. In S. Scalise et al. (eds.), *Universals of Language Today*, 165–184. Dordrecht: Springer.

Cinque, G. 2010. *The Syntax of Adjectives: A Comparative Study*. Cambridge, MA: MIT Press.

Cinque, G. 2014. *On the Movement Account of Greenberg's Universal 20: Refinements and Replies*. Ms., University of Venice.

Cinque, G. 2023. *Linearization: Toward a Restrictive Theory*. Cambridge, MA: MIT Press.

Costa, J., and N. Kula. 2008. Focus at the interface: Evidence from Romance and Bantu. In C. de Cat and K. Demuth (eds.), *The Bantu–Romance Connection*, 293–322. Amsterdam: John Benjamins.

Creider, C. 1989. *The Syntax of the Nilotic Languages*. Berlin: Dietrich Reimer Verlag.

Culbertson, J., and E. L. Newport. 2015. Harmonic biases in child learners: In support of language universals. *Cognition* 139: 71–82.

Culbertson, J., and E. L. Newport. 2017. Innovation of word order harmony across development. *Open Mind: Discoveries in Cognitive Science* 1: 91–100.

Culbertson, J., P. Smolensky, and G. Legendre. 2012. Learning biases predict a word order universal. *Cognition* 122: 306–329.

Culbertson, J., J. Franck, G. Braquet, M. Barrera Navarro, and I. Arnon. 2020. A learning bias for word order harmony: Evidence from speakers of non-harmonic languages. *Cognition* 204. https://doi.org/10.1016/j.cognition.2020.104392.

Den Dikken, M. 2006. *Relators and Linkers: The Syntax of Predication, Predicate Inversion, and Copulas*. Cambridge, MA: MIT Press.

Dik, S. 1983. Two constraints on relators and what they can do for us. In S. Dik (ed.), *Advances in Functional Grammar*, 267–298. Dordrecht: Foris.

Dik, S. 1997. *The Theory of Functional Grammar Part 1: The Structure of the Clause*. Berlin: Mouton de Gruyter.

Dryer, M. 1991. SVO languages and the OV:VO typology. *Journal of Linguistics* 27: 443–482.

Dryer, M. 2009. The branching direction theory of word order correlations revisited. In S. Scalise et al. (eds.), *Universals of Language Today*, 185–207. Dordrecht: Springer.

Dryer, M. 2012. On the position of interrogative phrases and the order of complementizer and clause. In T. Graf et al. (eds.), *Theories of Everything: In Honor of Ed Keenan* [UCLA Working Papers in Linguistics 17], 72–79. Los Angeles, CA: UCLA.

Dryer, M. 2018. On the order of demonstrative, numeral, adjective and noun. *Language* 94: 798–833.

Embick, D., and R. Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32: 555–595.

Ernst, T. 1994. M-command and precedence. *Linguistic Inquiry* 25: 327–335.

Frazier, L. 1980. Parsing and constraints on word order. *University of Massachusetts Occasional Papers in Linguistics* 6, article 9.

Giurgea, I. 2009. Adjective placement and linearization. In J. van Craenenbroeck (ed.), *Alternatives to Cartography*, 275–323. Berlin: Mouton de Gruyter.

Gorrell, P. 1995. *Syntax and Parsing*. Cambridge: Cambridge University Press.

Greenberg, J. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In J. Greenberg (ed.), *Universals of Human Language*, 73–113. Cambridge, MA: MIT Press.

Haider, H. 2005. How to turn German into Icelandic – and derive the OV–VO contrasts. *Journal of Comparative Germanic Linguistics* 8, 1–56.

Haider, H. 2014. *Symmetry Breaking in Syntax*. Cambridge: Cambridge University Press.

Halpern, A. 1995. *On the Morphology and the Placement of Clitics*. Stanford: CSLI Publications.

Hawkins, J. 1990. A parsing theory of word order universals. *Linguistic Inquiry* 21: 223–261.

Hawkins, J. 1994. *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.

Hawkins, J. 2008. An asymmetry between VO and OV languages: The ordering of obliques. In G. Corbett and M. Noonan (eds.), *Case and Grammatical Relations: Essays in Honor of Bernard Comrie*, 167–190. Amsterdam: John Benjamins.

Hawkins, J. 2014. *Cross-linguistic variation and efficiency*. Oxford: Oxford University Press.

Hein, J., and A. Murphy. 2022. VP-nominalization and the Final-over-Final Condition. *Linguistic Inquiry* 53: 337–370.

Hoekstra, T. 1984. *Transitivity: Grammatical Relations in GB-Theory*. Dordrecht: Foris.

Hyman, L. 2010. Focus marking in Aghem: Syntax or semantics? In I. Fiedler and A. Schwartz (eds.), *Expression of Information Structure: A Documentation of Its Diversity across Africa*, 95–116. Amsterdam: John Benjamins.

Jackendoff, R. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Jackendoff, R. 1990. On Larson's treatment of the double object construction. *Linguistic Inquiry* 21: 427–456.

Janke, V., and A. Neeleman. 2012. Ascending and descending VPs in English. *Linguistic Inquiry* 43: 151–190.

Kayne, R. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kayne, R. 2020. *Antisymmetry and Externalization*. Ms., New York University.

Kerstens, J. 1975. Over afgeleide structuur en de interpretatie van zinnen. Ms., University of Amsterdam.

Keyser, S. J. 1968. Review of Sven Jacobson, Adverbial Positions in English. *Language* 44: 357–374.

Knuth, D. 1968. *The Art of Computer Programming*, Vol. 1: Fundamental Algorithms. Reading, MA: Addison-Wesley.

Koster, J. 1974. Het werkwoord als spiegelcentrum. *Spektator* 3: 601–618.

Koster, J. 1978. Why subject sentences don't exist. In S. J. Keyser (ed.), *Recent Transformational Studies in European Languages*, 53–64. Cambridge, MA: MIT Press.

Langacker, R. W. 1969. On pronominalization and the chain of command. In D. A. Reibel and S. A. Schane (eds.), *Modern Studies in English*, 160–186. Englewood Cliffs, NJ: Prentice-Hall.

Larson, R. 1988. On the double-object construction. *Linguistic Inquiry* 19: 335–391.

Larson, R. 1989. *Light Predicate Raising* [Lexicon Project Working Papers 27]. Cambridge, MA: MIT Press.

Marantz, A. 1988. Clitics, morphological merger, and the mapping to phonological structure. In M. Hammond and M. Noonan (eds.), *Theoretical Morphology*, 253–270. San Diego, CA: Brill Academic Press.

Matushansky, O. 2002. *Movement of Degree/Degree of Movement*. PhD dissertation, MIT.

Medeiros, D. 2018. ULTRA: Universal Grammar as a universal parser. *Frontiers in Psychology* 9, article 155.

Moulton, K. 2013. Not moving clauses: Connectivity in clausal complements. *Syntax* 16: 250–291.

Moulton, K., and C.-H. Han. 2018. C-command vs. Scope: An experimental assessment of bound-variable pronouns. *Language* 94: 191–219.

Nchare, A. L. 2012. *The Grammar of Shupamem*. Doctoral dissertation, New York University.

Neeleman, A. 2017. PP-over-V meets Universal 20. *Journal of Comparative Germanic Linguistics* 20: 3–47.

Neeleman, A., and A. Payne. 2020. PP extraposition and the order of adverbials in English. *Linguistic Inquiry* 51: 471–520.

Neeleman, A., and H. Van de Koot. 2022a. The interpretation and distribution of temporal focus particles. *Natural Language & Linguistic Theory* 40: 793–835.

Neeleman, A., and H. Van de Koot. 2022b. PP extraposition and precedence. *Linguistic Inquiry* 53: 182–198.

Neeleman, A., and F. Weerman. 1999. *Flexible Syntax*. Dordrecht: Kluwer.

Nevins, A., and P. Weisser. 2018. Closest conjunct agreement. *Annual Review of Linguistics* 5: 219–241.

Nissenbaum, J. 2000. *Investigations of Covert Phrasal Movement*. Doctoral dissertation, MIT.

Overfelt, J. 2015. Unbounded successive-cyclic rightward DP-movement. *Lingua* 162: 1–31.

Partee, B., A. Ter Meulen, and R. Wall. 1990. *Mathematical Methods in Linguistics*. Dordrecht: Kluwer Academic.

Pearson, M. 2007. *Predicate Fronting and Constituent Order in Malagasy*. Ms., Reed College, Portland, OR.

Philip, J. 2012. *Subordinating and Coordinating Linkers*. PhD dissertation, UCL.

Philip, J. 2013. (Dis)harmony, the Head-Proximate Filter, and linkers. *Journal of Linguistics* 49: 165–213.

Phillips, C., and M. Wagers. 2007. Relating structure and time in linguistics and psycholinguistics. In G. Gaskell (ed.), *Oxford Handbook of Psycholinguistics*, 739–756. Oxford: Oxford University Press.

Rijkhoff, J. 1984. *The Principle of Head Proximity*. MA dissertation, University of Amsterdam.

Rizzi, L. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of Grammar*, 281–337. Dordrecht: Kluwer Academic.

Ross, J. R. 1969. On the cyclic nature of English pronominalization. In D. A. Reibel and S. A. Schane (eds.), *Modern Studies in English*, 187–200. Englewood Cliffs, NJ: Prentice-Hall.

Saito, M., and N. Fukui. 1998. Order in phrase structure and movement. *Linguistic Inquiry* 29: 439–474.

Shan, C.-C., and C. Barker. 2006. Explaining crossover and superiority as left-to-right evaluation. *Linguistics and Philosophy* 29: 91–134.

Sheehan, M. 2017. The Final-Over-Final Condition and processing. In Sheehan et al. (eds), 79–96.

Sheehan, M., T. Biberauer, A. Holmberg, and I. Roberts (eds.). 2017. *The Final-Over-Final Condition: A Syntactic Universal*. Cambridge, MA: MIT Press.

Steedman, M. 2020. A formal universal of natural language grammar. *Language* 96: 618–660.

Stowell, T. 1981. *Origins of Phrase Structure*. PhD dissertation, MIT.

Svenonius, P. 1994. The structural location of the attributive adjective. In E. Duncan et al. (eds.), *Proceedings of the 12th West Coast Conference on Formal Linguistics*, 439–454. Stanford, CA: CSLI Publications.

Takahashi, S. 2010. The hidden side of clausal complements. *Natural Language & Linguistic Theory* 28: 343–380.

Vanden Wyngaerd, G. 1989. Object shift as an A-movement rule. *MIT Working Papers in Linguistics* 11: 256–271.

Van der Wal, J. 2006. The disjoint verb form and an empty immediate after verb position in Makhuwa. In *Papers in Bantu Grammar and Description, ZASPiL 43*, 233–256.

Willer Gold, J., B. Arsenijević, M. Batinić, M. Becker, N. Čordalija, M. Kresić … and A. Nevins. 2018. When linearity prevails over hierarchy in syntax. *Proceedings of the National Academy of Sciences* 115: 495–500.

Williams, E. 1997. Blocking and anaphora. *Linguistic Inquiry* 28: 577–617.

Zeijlstra, H. 20222023. FOFC and what left-right asymmetries may tell us about syntactic structure building. *Journal of Linguistics* 59(1): 179–213.

Zwart, J.-W. 2009. Relevance of typology to Minimalist inquiry. *Lingua* 119: 1589–1606.