

# **DISPLACED DEPENDENT CONSTRUCTIONS**

Doug Arnold and Louisa Sadler  
University of Essex

Proceedings of the LFG13 Conference

Miriam Butt and Tracy Holloway King (Editors)

2013

CSLI Publications

<http://csli-publications.stanford.edu/>

## Abstract

An LFG treatment is proposed for ‘Displaced Dependents’ in English, including dependents of Degree Words, such as ‘so’, ‘too’, etc. and certain adjectives (e.g. ‘difficult’), as in examples like ‘too complex for anyone to understand’, ‘a difficult problem for anyone to understand’ where ‘for anyone to understand’ is dependent on ‘too’ and ‘difficult’ respectively, but is not adjacent to them.

## 1 Introduction

The specific topic of this paper is the analysis of ‘Displaced Dependents’ (DDs) in English, exemplified in (1); the term ‘Displaced Dependents’ is from Kay and Sag (2012),<sup>1</sup> who provide an impressive constructional analysis in the framework of Sign-Based Construction Grammar (SBCG). We attempt to provide an LFG analysis with comparable empirical coverage and theoretical appeal, without significant extension to the theoretical apparatus of the framework. A secondary, but broader, issue is comparison of the technical apparatus of LFG with Head-Driven Phrase Structure Grammar (HPSG) and SBCG: specifically, the discussion will bring out differences in the apparatus for lexical selection, and in the methods for controlling interactions among long-distance dependencies.

The DD construction occurs with degree words (*too*, *so*, *as*, *enough*, *more*, and synthetic comparatives); some degree denoting adverbs (e.g. *sufficiently*, as in (1f)); and some adjectives (e.g. *difficult*, *impossible*, *fun*), e.g. in (1g):

- (1) a. This problem is *too* complex (for anyone) to understand.
- b. This problem is *so* complex that we can't solve it.
- c. This problem is complex *enough* to be give you a headache.
- d. This problem is *more* complex than you can imagine.
- e. This problem is *harder* for them than you can imagine.
- f. The problems were *sufficiently* complex that we gave up.
- g. This is a *difficult* problem for anyone to solve without help.

Descriptively there are three key points to appreciate here. First, there is a genuine syntactic dependency. For example, *too* takes an optional non-finite clause with *for*, or a VP, but it is incompatible with a finite clause or a *than* clause:

- (2) a. This problem is *too* complex (for anyone) to understand.
- b. \*This problem is *too* complex that we can understand.
- c. \*This problem is *too* complex than we can understand.

Compare: *so* takes a *that* clause, as in (1b), and *more* and comparative adjectives take *than* clause dependents, as in (1c), (1d).

---

<sup>†</sup> We are grateful to several people for insightful comments and stimulating discussion, notably, Miriam Butt, Mary Dalrymple, Dag Haug, Tracy Holloway King, Joan Maling, Ida Toivonen, and Adam Przepiórkowski, as well as several anonymous referees, and other participants at LFG 2013 in Debrecen, Hungary. None of these people can be blamed for deficiencies in what follows.

<sup>1</sup>DDs are called ‘Indirect Complements’ in Huddleston and Pullum (2002, cf p549). Berman (1974) christened the construction the ‘Hard Nut’ construction (cf. the expression ‘a hard nut to crack’). Complements of degree words are sometimes referred to as ‘result clauses’, e.g. in Gueron and May (1984). Though we will concentrate exclusively on English, the construction appears in many other languages, including French and Polish (Adam Przepiórkowski, p.c.).

Second, it is a *discontinuous* dependency, and this discontinuity may be ‘optional’, or obligatory. In the case of adjectives like *difficult*, it is optional, in the sense of not being required in all circumstances (though still required in certain situations). For example, *difficult* does not require a dependent VP to be displaced, as witness the acceptability of (3) and (4). Displacement is allowed, as in (5), when the adjective is pre-nominal – and in fact in this situation it is required, as can be seen from (6).

- (3) A problem *difficult to solve without help* may be easier with help.
- (4) This problem is *difficult to solve without help*.
- (5) This is a *difficult problem to solve without help*.
- (6) \*This is a *difficult to solve without help* problem.

With degree words, displacement appears to be required – *too* is typical:

- (7) This problem is *too complex to understand*.
- (8) \*This problem is *too to understand* complex.
- (9) \*This problem is complex *too to understand*.

Dependent Displacement is not always optional or obligatory, of course. In fact, with most adjectives it is forbidden, as e.g. with *happy*, *grateful*, or *fond*:

- (10) a. The voters *happy with the election result* filled the streets.  
b. \*The *happy voters with the election result* filled the streets.
- (11) a. We got several letters from authors *grateful to the foundation for its support*.  
b. \*We got several letters from *grateful authors to the foundation for its support*.
- (12) a. No readers *fond of classic crime fiction* will want to miss this book.  
b. \*No *fond readers of classic crime fiction* will want to miss this book.

Finally, while (as will appear below) displacement can occur over some considerable distance, it is not unbounded. It is at least clause bounded – in (14) *too solve* has been displaced out to the complement of *seem*, and the result is unacceptable.

- (13) This will seem [to be a *difficult problem to solve*] when they encounter it.
- (14) \*This will seem [to be a *difficult problem*] when they encounter it *to solve*.

The remainder of the paper is structured as follows. In Section 2 we will briefly review two HPSG analyses and then introduce Kay and Sag (2012)’s SBCG analysis, which provides one of the leading ideas of our analysis (specifically, the idea that DDs involve a form of extraposition), and will also act as an empirical benchmark for our account. Our LFG treatment is presented in Section 3. Section 4 provides a conclusion.

## 2 Previous Analyses

We will not here attempt a comprehensive review of the literature on the DD construction. Instead, we will focus on three HPSG-style analyses that seem initially promising. The first two treat DDs as involving a form of ‘Predicate Composition’, the third treats DDs as involving a form of Extraposition.<sup>2</sup>

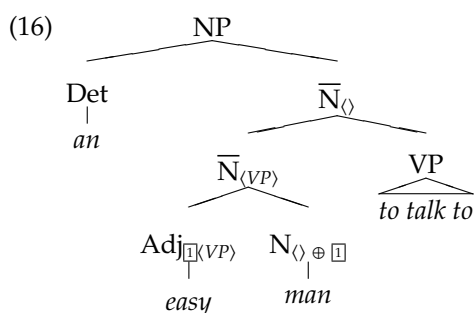
<sup>2</sup>Among the key sources this ignores are Bolinger (1972), Berman (1974), Huddleston and Pullum (2002), Chae (1992). As usual, Butt et al. (1999, see 119ff) contains relevant discussion. There is also a very extensive literature on comparatives – though this is mostly concerned with the internal properties of the

## 2.1 Flickinger and Nerbonne (1992)

In the context of a general discussion of lexical inheritance and adjectives, Flickinger and Nerbonne (1992) suggest an analysis of examples like (15) which involves some subcategorization requirements being ‘transferable’. In particular, adjectives like *easy* are specified as being allowed to transfer their subcategorization requirements to a sister (in this case the noun they modify).

(15) John is an easy man to talk to.

Thus, in (16), the *SUBCAT* requirement of *easy* (which, for the sake of discussion we take to be a list containing just a VP) is appended to the subcategorization requirement of *man* (which we assume is empty). This new list is passed up from the N to the mother  $\bar{N}$ , in the normal way. The VP requirement is satisfied by the presence of the actual VP *to talk to*, so that the subcategorization requirement of the highest  $\bar{N}$  is empty.<sup>3</sup>



In LFG, subcategorization restrictions are expressed in *PRED* values, so the spirit of this analysis could be captured in LFG as a form of syntactic complex predicate formation (e.g. Alsina, 1996; Butt, 1993). But this is not very appealing: the problem here is that in LFG complex predicate formation is standardly thought of as combining the *PRED* value (and hence the subcategorisation requirements) of a head with the *PRED* value of its argument. But here what we need is a method of combining the arguments of a head (e.g. a noun *man* or *problem*) with those of a *non*-argument – specifically, an *adjunct* (e.g. *easy*). Technical challenges aside, this is a theoretical extension which one would like to avoid, in the absence of independent motivation.

## 2.2 Kim and Sells (2011)

This particular problem is avoided in the analysis presented in Kim and Sells (2011). Here the degree word in a DD construction is treated as a ‘functor’ selecting (and successively combining with) an AP (e.g. *complex*), and a CP or VP dependent (*that we will never solve it*) – see (17) and (18).

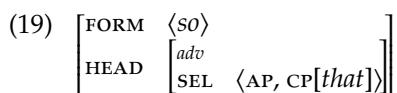
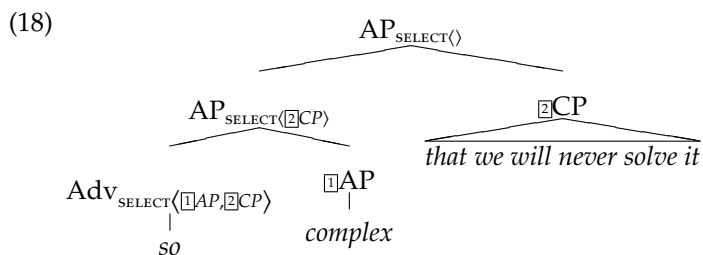
Standard HPSG (e.g. Pollard and Sag, 1994) contains a variety of mechanisms for lexical selection, Kim and Sells adopt a proposal from Allegranza (1998a), Van Eynde (2003) and van Eynde (2007), which restricts these to essentially two: selection of subjects and complements by heads, and selection of heads by ‘functors’ (e.g. adjuncts, and specifiers). Formally, this is implemented by having two phrase types

dependent clause, rather than the phenomenon of DDs.

<sup>3</sup>We write ‘ $\langle \rangle$ ’ for the empty list, and ‘ $\oplus$ ’ for append. Since  $\langle \rangle$  designates  $\langle VP \rangle$  on the Adj node, the specification ‘ $\langle \rangle \oplus \langle \rangle$ ’ which appears on the N is the same as the specification  $\langle VP \rangle$  that appears on the  $\bar{N}$ .

(*head-argument-phr*, and *head-functor-phr*), where in the latter, the non-head daughter bears a *SELECT* specification which must be satisfied by the head daughter. In a case like (17), this gives an analysis as in (18), based on a lexical entry for *so* as in (19).<sup>4</sup>

(17) This problem is *so complex that we will never solve it*.



The idea of (19) is that *so* 'SELECT's an AP, and a *that* marked CP. Hence in (18) it first combines with the AP *complex*, discharging the AP requirement, and forming the higher AP, this then combines with the *that* clause, discharging the requirement for a CP, leaving at the top an AP whose selection requirements are fully satisfied (*SELECT*  $\langle \rangle$ ).<sup>5</sup>

However, from an LFG point of view, the analysis is no more appealing than that of Flickinger and Nerbonne. The issue is that in LFG, as we already noted, the only general selection mechanism involves satisfaction of complementation requirements expressed in *PRED* values, and *PRED* values are borne by heads. Thus, we would have to assume that *so* is the head of *so complex*, and of *so complex that . . .*. But this runs counter to what one would naturally take to be the head-dependent relations, where the adjective *complex* is the head and *so* is a dependent (as Kim and Sells assume), and would lead one to expect the whole phrase to be an AdvP or DegP (depending on one's analysis of the category of *so*), when in fact it is an AP.<sup>6</sup>

### 2.3 Kay and Sag (2012)

<sup>4</sup>Kim and Sells do not discuss the DD construction directly, instead they focus on the so called 'Big Mess' construction (Berman, 1974), as in *so complex a problem that we will never solve it*, where *so complex . . . that . . .* is used attributively, modifying the nominal *a problem*. Their analysis is that in this construction *so* selects an AP (*big*), an NP (*a problem*), and a CP (*that . . .*). We think it is reasonable to assume that their analysis of the predicative use that we are concerned with simply omit selection of the NP.

Unrelated to this point, there is technical problem with Kim and Sells' analysis which seems to have passed unnoticed. In general, the idea is that head-functor structures involve a head discharging an element of the *SELECT* list of the non-head. Notice that at the top of (18) the CP discharges the *SELECT* value of the AP, requiring that the CP should be the head. This is of course entirely implausible (and incompatible with the category label AP on the topmost node).

<sup>5</sup>A first glance, it may be difficult to discern the difference between this approach and Flickinger and Nerbonne's, and indeed they are using the same basic formal machinery (lists, cancelation, percolation), but the grammatical principles involved are very different, as would be clear if the details were spelled out.

<sup>6</sup>Extending this approach to other DD constructions, one would have to assume that the adjective *difficult* is the head of *difficult problem to solve*, which one would normally assume is nominal, and that the adverb *sufficiently* is the head of *sufficiently complex that we gave up*, which is adjectival.

Kay and Sag (2012)'s analysis is more appealing from an LFG point of view. They analyse DD constructions as involving extraposition, analogous to extraposition of relative clauses as exemplified in (20). The account is formalised in the framework of Sign Based Construction Grammar (SBCG) (e.g. Sag, 2012).

(20) Then [ a girl — ] came in *who was clearly very pleased with herself*.

The key idea is that extraposition involves percolating an EXTRA list – that is, using similar apparatus as is used for handling long-distance dependencies (which are handled by percolating SLASH lists). Such dependencies are most easily understood as involving three components: a ‘bottom’, where the dependency is introduced, a ‘middle’, where the dependency is percolated, and a ‘top’ where the dependency is discharged.

If we consider an example like (21), the bottom of the dependency is the lexical entry for *so*, for which (22) gives the relevant details.

(21) (We were) *so* upset by the film *that we left early*.

(22) 
$$\left[ \begin{array}{l} \text{FORM} \langle so \rangle \\ \text{SYN} \left[ \begin{array}{l} \text{CAT} \quad [\text{SELECT} \quad [\text{SYN} \quad [\text{EXTRA} \quad \boxed{L} ]]] \\ \text{EXTRA} \quad \boxed{L} \oplus \langle S[\text{that}] \rangle \end{array} \right] \end{array} \right]$$

According to this, *so* is a functor (e.g. specifier or adjunct) of a head, and hence bears a SELECT feature describing this head. In the case of *so* the description includes the information that the head's EXTRA list is *L* (we ignore other restrictions, e.g. that the head should be adjectival). *So*'s own EXTRA list consists of *L* plus a *that*-marked S. In the typical case, e.g. *upset* in (21), the head's EXTRA list will be empty, so the EXTRA of *so* will just be  $\langle S[\text{that}] \rangle$ .

The ‘middle’ part of the dependency simply involves upwards inheritance of the EXTRA list from the functor daughter. The ‘top’ of the EXTRA dependency is handled by a special construction: the Head-Extraposition Construction (subtype of *headed-construction*), as in (23).

(23) 
$$hd\text{-extra}\text{-cxt} \Rightarrow \left[ \begin{array}{l} \text{MTR} \quad [\text{SYN} \quad \boxed{Y} ! [\text{EXTRA} \quad \boxed{L}]] \\ \text{DTRS} \quad \langle \boxed{H} : [\text{SYN} \quad \boxed{Y} [\text{EXTRA} \quad \langle \boxed{X} \rangle \oplus \boxed{L}]], \boxed{X} \rangle \\ \text{HD-DTR} \quad \boxed{H} \end{array} \right]$$

According to this, the top of an extraposition path consists of a mother, whose SYN value is  $\boxed{Y}$ , and two daughters: a head daughter whose SYN is also  $\boxed{Y}$ , and a non-head daughter  $\boxed{X}$ . The head daughter's EXTRA list consists of  $\boxed{X}$ , plus a (possibly empty) list of other extraposed elements  $\boxed{L}$  – which is also the EXTRA list of the mother.<sup>7</sup>

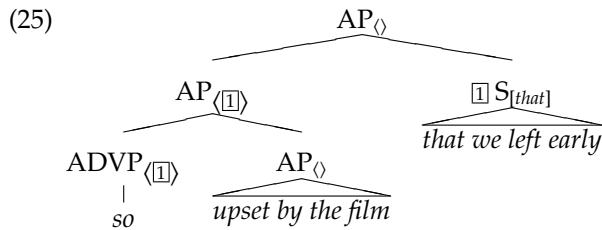
The central idea can be seen more clearly in (24), where the EXTRA list is represented as a subscript. Intuitively, the idea is to realise the first element of the EXTRA list as the right daughter of the head.

(24) 
$$\begin{array}{c} Y_{\langle \dots \rangle} \\ \swarrow \quad \searrow \\ Y_{\langle \boxed{X}, \dots \rangle} \quad X \end{array}$$

For example, in (25), the AP *upset about the film* has an empty EXTRA list, so the

<sup>7</sup>The ‘!’ here indicates that this is a default value. This is not important here.

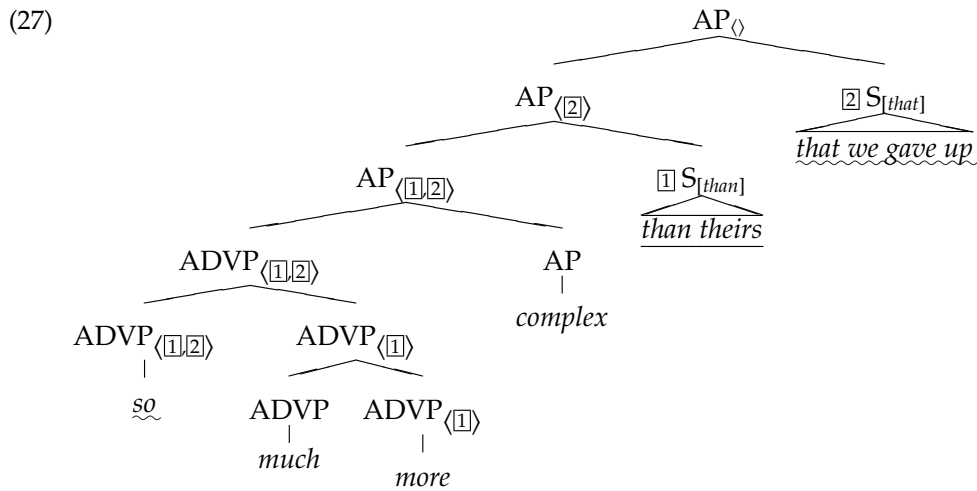
EXTRA list of the adverb *so* contains just the S[*that*] that is lexically specified. This specification is passed up to the AP (since the ADVP is the functor, i.e. non-head, daughter).<sup>8</sup> The top of the tree is an instance of the Head-Extrapolation Construction: the right daughter realises the first element in the EXTRA list of the head daughter, and the (empty) remainder of the list is passed upwards.



In an example like this it is not easy to see the motivation for using long-distance apparatus for this phenomenon. But the approach scales up easily to deal with cases involving multiple degree words, as in (26).

(26) (Our solution was) so much more complex than theirs that we gave up.

Here we see two degree words, both with displaced dependents: *so* . . . *that we gave up*, and *more* . . . *than theirs*. It is reasonable to assume that the lexical entry for *more* is like that for *so*, except that *more* requires its EXTRA list to contain a *than* marked S.



Again starting at the bottom, *more* passes its EXTRA requirement (the list containing the S[*than*]) to the adverbial *much more*.<sup>9</sup> *So* adds its EXTRA requirement (the list containing a S[*that*], here the 2) to the end of this, producing the list <math>\langle 1, 2 \rangle</math> and passes this to the

<sup>8</sup>We have labelled the pre-terminal nodes as phrasal (ADVP) because in HPSG and SBCG the distinction between X and XP is standardly taken to relate to complement saturation (phrases have empty COMPS list). On this interpretation, an adverb like *so* is in effect a lexical ADVP, and *complex* is an AP in this structure.

<sup>9</sup>Kay and Sag do not actually explain how this step works – notice that the EXTRA list should be inherited from the functor daughter, *much*, not as we have shown, from the head daughter, *more*. A possible solution would be to lexically specify *much* to share the EXTRA list of its head (i.e. like *so*, except that it does not add to the EXTRA list). Notice that all items that can modify items that permit extraposition will have to be similarly specified – e.g. adjectival modifiers like *very* which appear with DD adjectives, as in *very difficult* . . . *to please*. It is not clear if this creates problems.

mother ADV, and then to the A *so... complex*. The elements are then discharged from the EXTRA list, in order.

Notice that on this analysis, the displaced dependents of DD licensors like *so* and *more* are not complements: they originate on the EXTRA list of the licensors, and are never part of the complements list. This is presumably why the DDs never appear adjacent to such licensors – why displacement is obligatory: the only construction that would allow such adjacency would be the *head-complement* construction, which is obviously excluded if DDs are not complements.<sup>10</sup>

Cases where displacement is optional (e.g. with adjectives like *difficult*) can be handled by assuming the relevant dependents are complements when adjacent to the DD licensor, and that there is a non-branching lexical construction, which removes a complement from the VALENTS list (roughly the SUBCAT list of earlier work), and puts it on the EXTRA list:<sup>11</sup>

$$(28) \quad \begin{array}{c} \left[ \begin{array}{l} \text{word} \\ \text{SYN} \left[ \begin{array}{l} \text{VAL} \quad \langle NP \rangle \\ \text{EXTRA} \quad \langle PP_{of} \rangle \end{array} \right] \end{array} \right] \\ | \\ \left[ \begin{array}{l} \text{word} \\ \text{SYN} \left[ \begin{array}{l} \text{VAL} \quad \langle NP, PP_{of} \rangle \\ \text{EXTRA} \quad \langle \rangle \end{array} \right] \end{array} \right] \end{array}$$

(29) They are not very *proud* always *of their real achievements*.

An interesting feature of the analysis is that elements are removed from the EXTRA list in strict order, from the front. In the case of (26) above, this gives the right result, in that it ensures that the dependencies between DD licensors and DDs are nested. Thus (26) is good, whereas (30), which shows a crossing dependency, is not:

(30) \*Our solution was *so* much *more* complex *that we gave up* than theirs.

However, Kay and Sag observe that not all DD licensors are alike in this respect. In particular, some DD licensors, like *more*, allow *crossing* dependencies. Alongside (31) with nested dependencies, we have (32) where the dependencies cross.

(31) This was *more* difficult in those days *for people to solve* *than we can imagine*.

(32) This was *more* difficult in those days *than we can imagine* *for people to solve*.

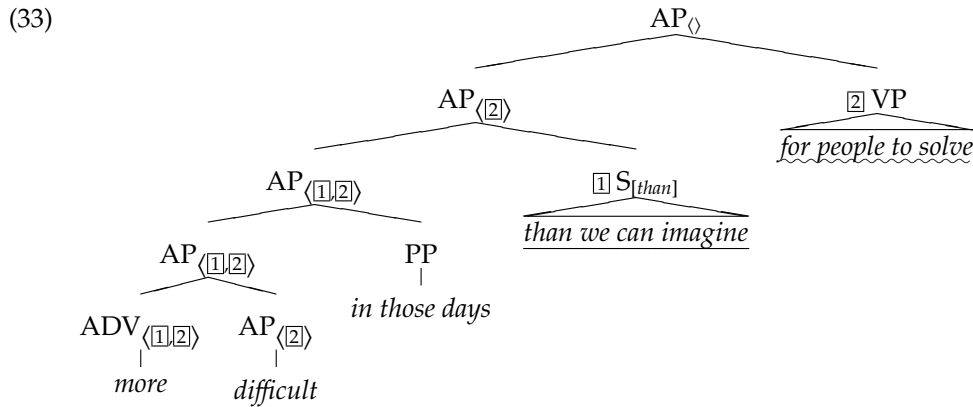
Kay and Sag's account of this is straightforward: whereas items like *so* add their complements to the end of their head's EXTRA list, lexical items like *more* are allowed to scramble their complements into their head's EXTRA list. In the case of (31) the *than* dependent should be added to the end of the EXTRA list of *difficult*, in the case of (32) it should be scrambled in before the EXTRA list of *difficult*. (33) gives the representation

<sup>10</sup>This is not completely accurate: it is not clear what prevents examples like \**so that we gave up interesting*, as a result of *so* to speak 'string vacuous' extraposition. On the face of it, there is nothing to stop *so* and *that we gave up* combining in a head-extraposition construction, which then functions as an adjunct of *interesting*. The same problem will arise with DD adjectives. Notice it will not do to require the head in a head-extraposition structure to be phrasal, because as noted above (footnote 8) on the standard interpretation of this in HPSG and SBCG words like *so* are inherently phrasal.

<sup>11</sup>However, Kay and Sag do not offer an account of why items like *difficult*, which allow displacement in general, *require* it in some circumstances, as in \**a difficult for us to solve problem*. We will address this problem in our LFG treatment below.



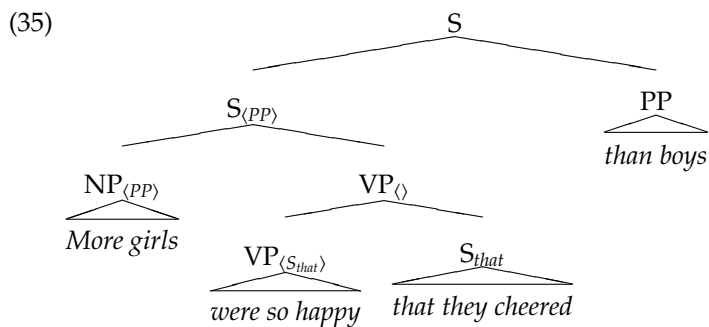
of (32).



As regards the boundedness of Extraposition, Kay and Sag note the following:

- (34) a. **More** girls were so happy that they cheered **than boys**.  
 b. \***More** girls were so happy **than boys** that they cheered.

At first glance, this is puzzling, because we observed that *than* dependents (dependents of *more*) could shuffle in among other displaced dependents, but in (34b) it appears this is not allowed. Kay and Sag suggest this is a bounding effect. We are dealing here with two distinct EXTRA lists: one from the subject (*more girls...*) and one from the predicate (*so happy...*). The contrast in (34) can be explained if the latter has to be realised before the former, for example if the *that* clause is realised as a sister of VP, and the *than* clause is realised as a sister of S.



To achieve this, Kay and Sag suggest that the subject-predicate construction has a constraint to the effect that the EXTRA list on a finite VP must be empty (so finite VP acts as a barrier to extraposition). Notice that (35) satisfies this constraint, but that it cannot be satisfied in a representation of (34b).<sup>12</sup>

Kay and Sag provide an account of the semantics of DD constructions, formulated in Minimal Recursion Semantics (MRS). The details of this are not important here. The key points are (a) that the non-branching constructions that move complements to the EXTRA list do not disturb the semantics; (b) *so*, *too*, etc. are interpreted as generalized

<sup>12</sup>It is not clear that this is the only, or the best account possible. Kay and Sag do not say how EXTRA lists percolate in head-complement or subject-predicate constructions, though they do hint (p248) at inheritance from the first daughter being the norm.

quantifiers which bind a degree variable and have a restriction associated with the expression they modify, and a scope associated with the element they add to the EXTRA list (i.e. the DD). That is, in either case, the semantics is handled ‘locally’ within the lexical item: displacement does not affect it. For example (36) has the roughly the semantics in (37).<sup>13</sup>

(36) Kim was so happy that she cried.

(37)  $so(\delta, happy(s_0, Kim, \delta), cried(s_1, Kim))$

To sum up: Kay and Sag give an account of the syntax and semantics of DD constructions. Displacement is treated as a form of extraposition, and extraposition is handled by a collection of constructions. In particular, a construction that realises elements of the EXTRA list as right daughter to a head. Displacement possibilities are reflected in other constructions also, because of the way that EXTRA lists are percolated (or not, cf. the way the boundedness of the displacement is encoded by having constructions require certain EXTRA lists to be empty). There are some lexical stipulations, e.g. the semantics is given lexically, and the cases where displacement is obligatory (specifically, with degree words like *so*, *too*, etc.) arise because of the lexical properties of those items: they have no complements, their dependents appear on their EXTRA lists. Cases where displacement is optional arise where items are able to occur in the construction that moves complements from the COMPS list to the EXTRA list. Clearly, this will have to be stated in the lexicon. Likewise, cases where displacement is forbidden will arise because the relevant items are unable to appear in these constructions. Again, this will have to be stated in the lexicon.

The key idea here is that DDs involve a non-local dependency. In the next section we will explore how this idea can be developed to provide an account of the phenomena that can be expressed without significant extension to the theoretical apparatus of LFG.

### 3 LFG Analysis

The most obvious and natural way to capture the syntactic dependency that exists between DDs and their licensors is to treat DDs as complements (that is, COMPS or XCOMPS) of their licensors.<sup>14</sup> The question then arises whether DDs have, in addition to their in-clause function, an additional (e.g. discourse) function in their displaced (‘upstairs’) location: that is, whether there is an EXTRA feature, analogous to TOPIC, FOCUS, etc. We are not aware of any well developed LFG accounts of extraposition that would bear on this issue, or any clear evidence one way or another, but the simplest, most theoretically parsimonious, assumption is that they do not. Hence, we will assume that extraposition is a purely c-structure phenomenon. For simplicity, we will also assume that degree words and attributive adjectives are adjuncts of their heads.

That is, we assume an example like (38) receives a representation like (39).<sup>15</sup>

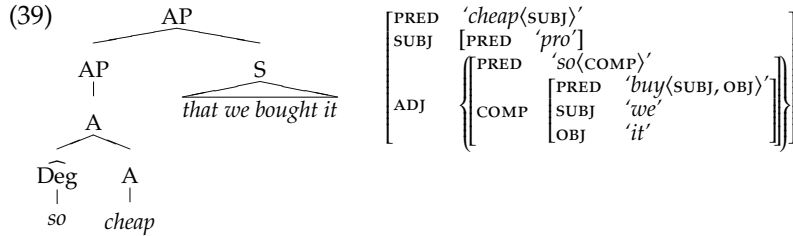
---

<sup>13</sup>Here  $\delta$  is the degree variable associated with adjectives like *happy*,  $s_0$  and  $s_1$  are situation variables: the meaning is roughly there is a degree  $\delta$  such that Kim was happy to degree  $\delta$  in situation  $s_0$ , and this has a consequence that Kim cried in situation  $s_1$ .

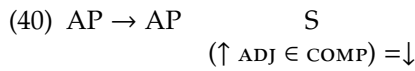
<sup>14</sup>For simplicity, we will assume here that even complements that lack overt subjects (i.e. that are superficially VPs) are COMPS.

<sup>15</sup>We assume that the copula does not contribute a PRED value here, but nothing hangs on this. We will

(38) It was *so cheap* *that we bought it*.

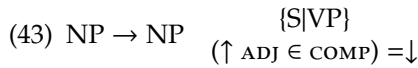
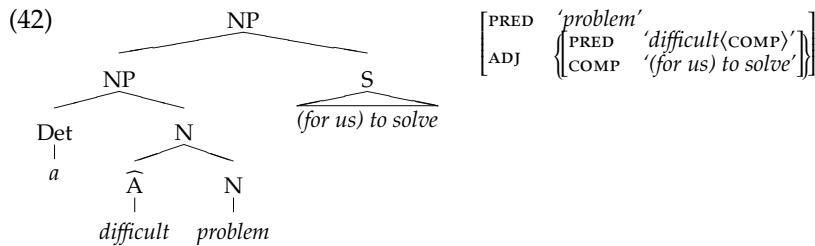


Producing the c- and f-structures in (39) will require a phrase-structure rule like (40), which makes the f-structure of the S into the COMP of an element of the ADJUNCT set of the AP.



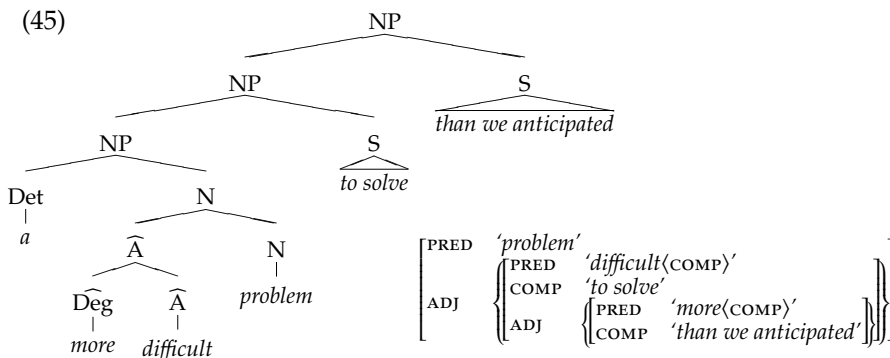
Similarly (41) will receive a representation like (42), involving a rule like (43).

(41) This is a *difficult* *problem* (*for us*) *to solve*.



The constraints in these rules must be generalised to deal with examples of longer distance displacement, as in (44), where *than we expected* is not a COMP of an adjunct of the head of NP, but of an adjunct's adjunct: see (45).

(44) This is a more difficult problem to solve than we anticipated.



Generalising the above, we arrive at (46), which says that a phrase final S or VP explain the significance of the 'hat' on  $\widehat{\text{Deg}}$  here, and on  $\widehat{A}$  in (42), below.

can be the COMP of a head, or of an adjunct of head, or of adjunct of an adjunct of the head, etc.

$$(46) \text{XP} \rightarrow \text{XP} \quad \left( \begin{array}{c} \{S|VP\} \\ (\uparrow (\text{ADJ} \in)^* \text{COMP}) = \downarrow \end{array} \right)$$

Notice that, given this, semantics along the lines suggested by Kay and Sag will be entirely straightforward. DD adjectives will be like other adjectives that take complements – when used attributively they will be associated with meaning constructors that consume the resources contributed by their complements, and produce constructors like those for simple adjectives (see Dalrymple (2001, p260ff)). Degree words will correspond to constructors that consume the resources associated with their complements, and the adjectives they modify, and produce resources with meanings like that given by Kay and Sag above. Roughly, the glue expression required is as in (47) (intuitively, this produces the interpretation of *too Adj VP* by consuming in turn the semantics of the DVAR (degree variable) of an item of which *too* is an adjunct ((ADJ ∈ ↑)<sub>σ</sub>DVAR), the item of which *too* is an adjunct, and the complement of *too*):

$$(47) \lambda d.\lambda P.\lambda Q.\text{too}(d, P(d), Q) : ((\text{ADJ} \in \uparrow)_{\sigma} \text{DVAR}) \multimap (\text{ADJ} \in \uparrow)_{\sigma} \multimap (\uparrow \text{COMP})_{\sigma} \multimap \uparrow_{\sigma}$$

The schema in (46) is sufficient to license DDs, but it over-generates. As we observed above, displacement of dependents is sometimes necessary (e.g. with *too*, as in (48), (49)), sometimes optional (as with *difficult*, in cases like (50) and (51)), and sometimes forbidden (as with ‘normal’ adjectives like *grateful*, in (53), as compared to (54)), and *difficult* in an example like (52)).

- (48) This problem is *too* complex to understand.
- (49) \*This problem is *too* to understand complex.
- (50) This is a *difficult* problem to solve without help.
- (51) A problem *difficult* to solve without help may be easier with help.
- (52) \*A *difficult* to solve without help problem
- (53) \*She is a *grateful* person to her parents.
- (54) She is a person *grateful* to her parents.

The key assumption in our account of this will be that certain phrase structure positions – specifically, pre-adjectival and pre-nominal positions – are defective in not allowing the full range of phrasal projections of categories. This is not a novel suggestion. For example, Poser (1992) introduced the idea of ‘small’ categories, which was taken up by Sadler and Arnold (1994); Abeillé and Godard (2000)’s notion of ‘LITE’ expressions implements a similar idea in HPSG; Toivonen (2003) introduced the concept of ‘non-projecting’ categories. We will adopt a slightly modified version of Toivonen’s idea.

According to Toivonen, (*loc cit* p55ff), as well as the usual X-bar categories ( $X^0$ ,  $\bar{X}$ , XP), there are ‘non-projecting’ categories ( $\widehat{X}$ ), which do not license higher levels of phrase structure ( $X^0$  and  $\widehat{X}$  are otherwise identical: intuitively, a  $\widehat{X}$  is just an  $X^0$  that cannot appear in an XP, though it may of course have a different distribution because of this, e.g. it will not be able to combine with complements and adjuncts in the normal way, and may appear in different c-structure positions). As well as the usual schemata in (55), which license structure in projections of  $X^0$  and  $\bar{X}$ , Toivonen has the schemata in (56) which allow adjunction to of YP to XP, and  $\widehat{Y}$  to  $X^0$ .<sup>16</sup>

<sup>16</sup>We make a minor departure from Toivonen in (55) and (56): where we have YP, she has YP\*. Thus,

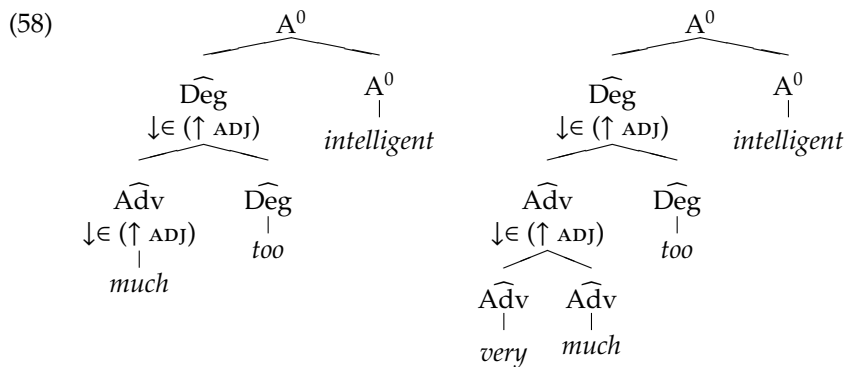
- (55) a.  $XP \rightarrow \bar{X} \ YP$   
 b.  $\bar{X} \rightarrow X^0 \ YP$

- (56) a.  $XP \rightarrow XP \ YP$  (Adjunction to  $XP$ )  
 b.  $X^0 \rightarrow X^0 \ \bar{Y}$  (Adjunction to  $X^0$ )

To these we will add the schema in (57), which allows modification of non-projecting categories (by non-projecting categories).<sup>17</sup>

- (57)  $\bar{X} \rightarrow \bar{X} \ \bar{Y}$

We use the possibility of modification of non-projecting categories to provide an analysis of examples like *much too (intelligent)*, *very much too (intelligent)* which we assume have structures as in (58) (with *much* and *very much* being adjuncts of *too*).<sup>18</sup>



We make the following rather natural assumption:

- (59) In general (i.e. by default), lexical items in categories that allow complements are  $X^0$ ;  $\bar{X}$  items do not have complements.<sup>19</sup>

The effect of this is just to establish  $\bar{X}$  as the marked case for lexical items, as compared to  $X^0$ . This is a very natural assumption: what it says in effect is that lexical items that might in principle have complements (in virtue of their category) will appear in phrase structure contexts where those complements could be realised.

We require two additional assumptions that are specific to English:

- (60) pre-head (in particular, pre-nominal and pre-adjectival) positions are restricted to  $\bar{X}$ s (e.g.  $\bar{A}$ ,  $\bar{Adv}$  and  $\bar{Deg}$ );  
 (61) post-head (esp. post-nominal and predicative positions) are restricted to XPs (i.e. projections of  $X^0$ )

We will now discuss some implications of these assumptions, and show that they

we generally assume a recursive phrase structure, where she assumes a flatter structure. We do not think anything hangs on this.

<sup>17</sup>This is less restrictive than Toivonen's position – she in fact claims that the schemata in (55) and (56) are the only ones that are permitted. However, we do not think the addition causes empirical problems for her proposals.

<sup>18</sup>Here and other representations, we generally suppress the  $\bar{A}$  between  $A^0$  and  $AP$ , for reasons of space.

<sup>19</sup>Here, we assume that this means their  $PRED$  list is empty; that is, we are taking 'complement' to cover all grammatical functions, including  $SUBJ$ . Consequently, we assume that attributive adjectives do not select  $SUBJ$  functions. This assumption could be discarded by an appropriate re-definition of 'complement' as 'GF-( $SUBJ$ )'

account for the range of data we require.

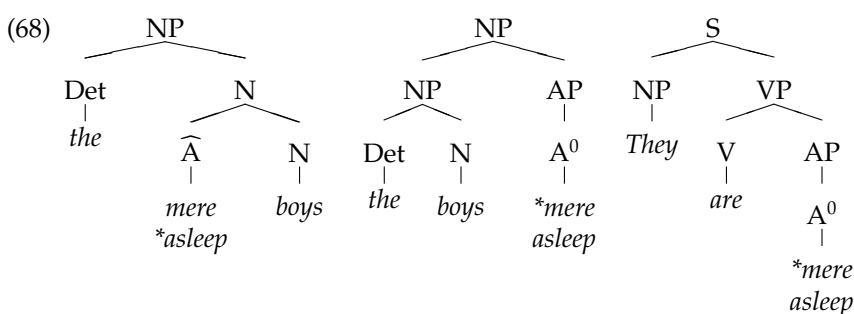
First, consider the fact that certain adjectives appear only pre-nominally or only post-nominally. Examples of adjectives the first kind include *mere*, *former*, *alleged*, *sheer*. These adjectives cannot appear post-nominally, or be used predicatively:

- (62) They are mere boys.
- (63) \*They are boys mere.
- (64) \*They are mere.

This follows, given our assumptions, if these adjectives are lexically specified as  $\widehat{A}$ . They are permitted to appear pre-nominally, but the only way they could appear post-nominally or predicatively would be if they could project an AP, which would require them to be  $A^0$ s. See the structures in (68).

Conversely, for adjectives like *awake*, *asleep*, *alive*, which cannot appear pre-nominally, but can appear post-nominally and predicatively: this behaviour follows if they are only  $A^0$ s. See (68).

- (65) \*The asleep boys were awakened by the noise.
- (66) The boys asleep were awakened by the noise.
- (67) The boys were asleep.



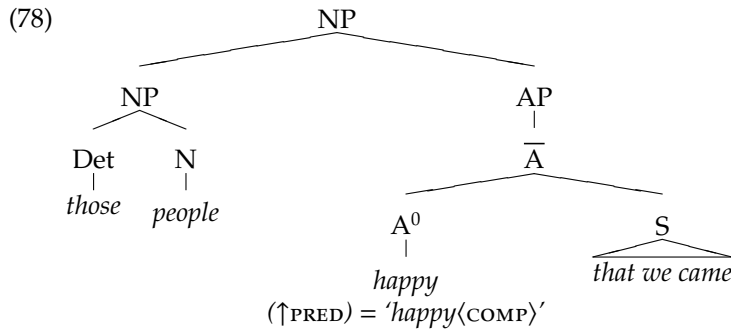
Secondly, consider the case of ‘normal’ adjectives (*happy*, *grateful*, *fond*, etc.), which appear post-nominally, and predicatively, with or without complements and post-nominal modifiers (as in (69)–(73)), but which can only appear pre-nominally when they do not have complements and are not themselves post-modified (as in (74)–(77)), even if as in (77) their dependents are displaced.

- (69) those people *happy that we came*
- (70) They are *happy that we came*.
- (71) those people *happy at the moment*
- (72) They are *happy at the moment*.
- (73) Those people *happy* (irritate me, though I can tolerate them when they are sad).
- (74) those *happy* people
- (75) \*those *happy that we came* people
- (76) \*those *happy at the moment* people
- (77) \*those *happy* people *that we came*

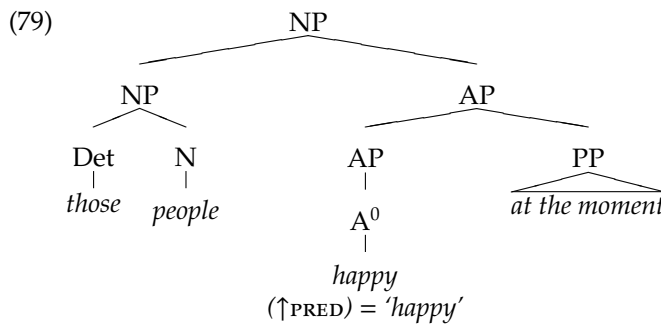
Our account is that such ‘normal’ adjectives lead double lives as  $\widehat{A}$ s and  $A^0$ s, conforming to (59). That is, they are  $A^0$ s by default (since A is a category that allows complements), but can also appear as  $\widehat{A}$ s – though in this case they will not project

any higher phrase structure. When they have complements, e.g. in the use of *happy* that takes a sentential complement (*'happy<COMP>'*) they will of necessity appear in APs, and similarly when they are post-modified.

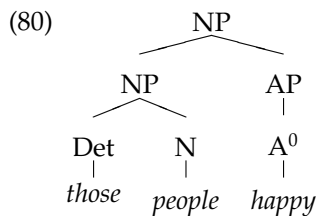
Consider an example like (69), where *happy* has a sentential complement *that we came*. This will have a structure like (78): X-bar principles (cf the schema in (55b)) dictate that the S complement has to be a sister to  $A^0$ , and since we have an  $A^0$ , we also get an AP, which is permitted post-nominally (and predicatively as in (70)).<sup>20</sup>



Similarly, an example like (71), where there is a post-nominal PP modifier, will have a representation like (79). X-bar principles dictate that the PP modifier *at the moment* has to be adjoined to AP, and AP requires an  $A^0$  to project the structure (cf the X-bar schema in (56)).



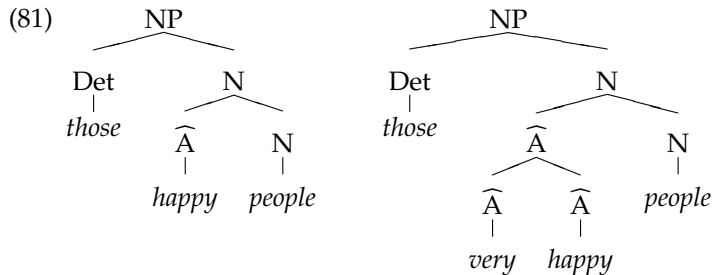
An example like (74) (*those people happy*) will have a structure like *the boys asleep* in (68): simply because an item is an  $A^0$  does not mean it has to have a complement, and one can have an AP without there being any post-modifiers.



A normal pre-nominal adjective, such as in (74), will be an  $\widehat{A}$ , giving a represen-

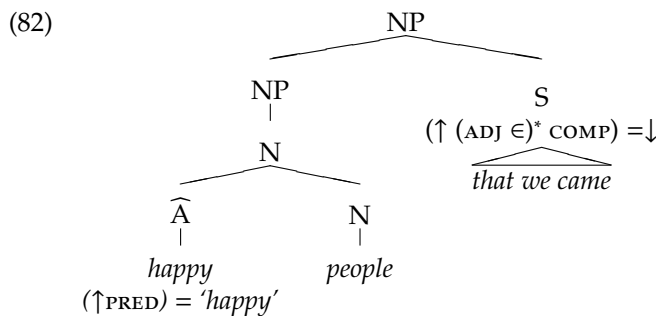
<sup>20</sup>Notice that in (78) or the following cases, a simple  $A^0$  can be replaced by a pre-modified  $A^0$  with a structure like the following:  $[_{A^0} [_{\widehat{A}^0} \text{very}] ]_{A^0}$  *happy* ], as permitted by the schema in (56b).

tation like the following. The X-bar schema in (57) permits the  $\widehat{A}$  to be pre-modified, as in the second structure in (81).



The ungrammaticality of examples like (75), and (76) is straightforward: it follows from the fact that an expression like *happy that we came* or *happy at the moment* must be an AP, and pre-nominal position is restricted to  $\widehat{A}$ s, excluding APs.

The ungrammaticality of (77), where *happy* appears in a DD construction, is equally straightforward: when *happy* is pre-nominal, it must be an  $\widehat{A}$ ; but this in general means it cannot have complements (in conformity to our assumption in (59)), hence it is automatic that it cannot have *displaced* complements. The representation would be along the lines of (82), which is a coherence violation (*that we came* is required to be a complement inside an adjunct somewhere lower in the structure – but there is no PRED that allows this).



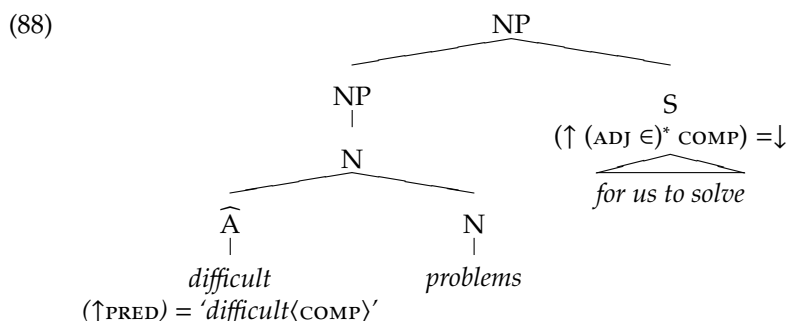
Thirdly, consider the case of adjectives, like *difficult*, that permit, but do not always require, DDs. These are exceptions to (59). They also lead a double life as  $\widehat{A}$  and a  $A^0$ , but in both cases they can have complements. But of course, if they are  $\widehat{A}$  (e.g. pre-nominal) their complements cannot be realised locally – they have to be displaced. Thus, we get the following data. The predicative use in (83) and the post-nominal use in (84) are exactly like the corresponding cases with *happy*, as is the contrast between the pre-nominal case with and without a complement in (85) and (86). The interesting case is (87).

- (83) These problems are difficult for us to solve.
- (84) Problems difficult for us to solve alone (can be easier with help).
- (85) \*A difficult for us to solve problem has been identified today.
- (86) A difficult problem has been identified today.
- (87) *difficult* problems for us to solve

The representation of (87) will be as in (88). This is parallel to (82), except that here



the equation on the extraposed complement can be satisfied, making *that we came* the complement of *difficult*, and avoiding the coherence violation.



Finally, we come to degree words (*so, too, more, etc.*), whose complements displace obligatorily. They resemble adjectives like *difficult* in being exceptions to (59). The difference is that they lead a *single* life as Deg.

- (89) *too difficult for us to solve*
- (90) \**too for us to solve difficult*
- (91) \**difficult too for us to solve*

Examples like (89) are allowed in the same way as examples like *difficult problems for us to solve* (see above). Examples like (90) are excluded because *for us to solve* is a complement, and must attach as sister to a Deg<sup>0</sup>, but *too* is a Deg. Examples like (91) have two potential analyses. In one *for us to solve* would be a complement sister of *too*. This is excluded for the same reason as (90). An alternative (but equally non-viable) analysis would be that *for us to solve* is an extraposed complement of *too* (via an application of ‘string vacuous’ extraposition). However, this would still leave *too* in a post-head position, which would require it to be the head of DegP, which in turn would require it to be a Deg<sup>0</sup>, which we have excluded.

The essence of our analysis is that degree words reflect an interesting mismatch between what one might think of as the demands of different components of the grammatical system. Lexically, they require complements, and this is reflected in f-structure, but their c-structure category is such that they can never realise their complements locally – hence for them extraposition is obligatory.

## 4 Conclusion

The original goal of the paper was to provide an LFG analysis of DDs that has comparable empirical coverage and theoretical appeal to existing accounts, specifically that of Kay and Sag (2012). With one reservation which we will address below, we think this goal has been achieved. The account we have given has a good empirical fit for the data, and while it is inspired by the account in Kay and Sag (2012), in that it treats DDs as involving extraposition, it uses only existing LFG machinery; specifically: (i) a novel, but very simple, story about extraposition (cf. (46)); (ii) a slight extension of an independent theory of non-projecting categories and their distribution (based on Toivonen (2003)); and (iii) plausible assumptions about lexical properties of Degree Words and DD adjectives: essentially that they show a mismatch between complementation and phrase structure projection properties – in being  $\widehat{X}$  items that take

complements.

There remain, however, a number of problems and open questions, which we will briefly review.

On the empirical side, our account under-generates in at least one respect. A descriptive consequence of our analysis is that pre-nominal adjectives cannot have complements. This generalisation is well-known to be false for at least two classes of case. One is exemplified by the pre-nominal adjective in (92). We do not find this very troubling, because it is part of a much more general phenomenon which allows almost any phrase to be converted into a pre-nominal modifier, cf. for example the clause in (93). These expressions require special intonation, and seem to require special treatment independent of the phenomena we have looked at here.

- (92) She's one of those very-tall-for-their-age-and-not-very-happy-with-it little girls.  
(93) I'm having one of those I'm-so-sick-of-this-bloody-job-that-I-could-scream days.

A more problematic collection of counter-examples is given in the following.

- (94) a taller than average child  
(95) a difficult to please person  
(96) a hard to solve problem

We have no good account of these, which are problematic for all other accounts, including those reviewed above. However, we note that not only are there restrictions on the adjectives that permit this (e.g. *\*a happy to be considered interviewee*), there are also restrictions on the kind of NP that can appear inside the construction. Thus, while *average* is acceptable in (94), it cannot be replaced with a referential or quantificational noun:

- (97) \*a taller than me child (cf. a child taller than me)  
(98) \*a taller than this mark child (cf. a child taller than this mark)  
(99) \*a more expensive than most houses car (cf. a car more expensive than most houses)

The account also over-generates. Our extraposition rule allows extraposition of COMP and from inside adjuncts; this is over-general. For example, relative clauses are adjuncts, and can contain complements, as in (100), cf. the representation in (102) but these complements cannot be extracted, as (101) shows.

- (100) the man [ who Kim thinks [that Sam admires] ] in the brown suit  
(101) \*the man [ who Kim thinks \_\_\_ ] in the brown suit [that Sam admires]

- (102) 
$$\left[ \begin{array}{l} \text{PRED} \quad \textit{man} \\ \text{ADJ} \quad \left\{ \begin{array}{l} \text{PRED} \quad \textit{'thinks'} \\ \text{SUBJ} \quad \textit{'Kim'} \\ \text{COMP} \quad \textit{'that Sam admires'} \end{array} \right\} \end{array} \right]$$

Another area where our account is potentially deficient relates to the issue of boundedness. Recall that Kay and Sag treated the contrast between (34a) and (34b), repeated here, by requiring that the EXTRA list of finite VP be empty.

- (103) a. **More** girls were so happy that they cheered **than boys**.  
b. \***More** girls were so happy **than boys** that they cheered.

This is a natural thing to state in HPSG or SBCG, where the S and VP are associated with distinct data-structures. But it is *not* an easy thing to state in LFG, where there is no simple LFG equivalent of ‘finite VP’ as opposed to ‘finite S’ (finiteness is an f-structure property, and S and VP have the same f-structure). It would, of course, be possible to express this constraint by means of a complex c-structure category, but one would prefer to have independent motivation for such a move.

Fortunately, it seems this is not the correct characterisation. In particular, notice that finiteness is not the issue, since the same contrast arises with non-finite VPs, such as the non-finite complement of *prefer* (*prefer that . . . be . . .*):

- (104) a. I would prefer that **more** girls be so happy that they cheer **than boys** than the reverse.  
 b. \*I would prefer that **more** girls be so happy **than boys** that they cheer than the reverse.

However, an alternative account is possible: the facts follow if extraposition to S is limited to the elements of the subject, and extraposition to VP from subject is excluded. That is, if the constraints on S and VP are as follows:

- (105)  $S \rightarrow S \quad \begin{matrix} \{S|VP\} \\ (\uparrow \text{SUBJ} (\text{ADJ } \in)^* \text{COMP}) = \downarrow \end{matrix}$   
 (106)  $VP \rightarrow V \quad \begin{matrix} \{S|VP\} \\ (\uparrow (\text{GF} - \text{SUBJ}) (\text{ADJ } \in)^* \text{COMP}) = \downarrow \end{matrix}$

However, this is certainly not the whole story. There are gaps in Kay and Sag’s account here as well. For example, while complements of *more* can be displaced from subjects, not all DDs can move so far. For example, complements of DD adjectives like *difficult* cannot:

- (107) A difficult problem *to solve* was raised at the meeting.  
 (108) \*A difficult problem was raised at the meeting *to solve*.

We leave this as a topic for further research.

Comparison with the HPSG based account discussed in Section 2 brings out the very limited apparatus LFG has for lexical selection when compared to HPSG: there is nothing in LFG comparable to the *MOD*, or *SPEC* features, or the more recent *SELECT* feature that HPSG analyses can use to allow non-heads to select the sorts of head they will accompany. Given that we have provided an analysis that manages without such apparatus, this does not seem to be a disadvantage.

But there is one respect in which our account does not match that of Kay and Sag (2012), and this is a reflection of a genuine and important difference between the theoretical apparatus of HPSG/SBCG and LFG. Our account of DDs, like that of Kay and Sag, involves extraposition, and our account of extraposition uses functional uncertainty equations, which are in their nature *independent* statements about relationships. This is adequate for the cases above (with *more*), which allow scrambling of DDs from different licensors. But Kay and Sag pointed out that some DD licensors (e.g. *so*) require nesting, e.g. (30) and (31) above, repeated here:

- (109) (The problem is) so difficult to solve that we gave up.  
 (110) \*(The problem is) so difficult that we gave up to solve.

This is handled straightforwardly, albeit by stipulation, in Kay and Sag's account: elements are discharged from the EXTRA list in strict order (from the front), items like *so* push their complements onto the end of the EXTRA list; hence these dependencies nest. Modelling long-distance dependencies with lists as in HPSG and SBCG gives very fine control over relations of the same kind. It is not clear how this can be expressed nicely in a functional uncertainty account: to get the same level of control in a functional uncertainty based account would appear to require a variety of off-path constraints and/or otherwise unmotivated features, which is not very appealing (cf. Dalrymple and King (2013) for discussion of phenomena that raise similar issues).

## References

- Anne Abeillé and Danièle Godard. The syntactic flexibility of adverbs: French degree adverbs. In Stefan Müller, editor, *Proceedings of the HPSG-2003 Conference, Michigan State University, East Lansing*, pages 26–46, Stanford, 2003. CSLI Publications.
- Anne Abeillé and Danièle Godard. French word order and lexical weight. In R.D. Borsley, editor, *The Nature and Function of Syntactic Categories*, Syntax and Semantics. Academic Press, 2000. ISBN 9780126135329.
- Valerio Allegranza. Determiners as functors: NP structure in Italian. In Sergio Balari and Luca Dini, editors, *Romance in Head-driven Phrase Structure Grammar*, volume 75 of *CSLI Lecture Notes*, pages 55–108. CSLI Publications, Stanford, 1998a.
- Alex Alsina. *The role of argument structure in grammar: Evidence from Romance*. CSLI Publications, Stanford, CA, 1996.
- Doug Arnold and Louisa Sadler. Noun-modifying adjectives in HPSG. Technical report, Department of Language and Linguistics, University of Essex, April 1992.
- Arlene Berman. *Adjectives and Adjective Complement Constructions in English*. PhD thesis, Dept of Linguistics, Harvard University, 1974.
- Dwight Bolinger. *Degree words*. Mouton, The Hague, 1972.
- Miriam Butt. *The structure of complex predicates in Urdu*. PhD thesis, Stanford University, Department of Linguistics, 1993.
- Miriam Butt, María-Eugenia Niño, and Frédérique Segond. *A Grammar Writer's Cookbook*. CSLI Publications, Stanford, CA, 1999.
- Hee-Rahk Chae. *Lexically Triggered Unbounded Discontinuities in English: an Indexed Phrase Structure Grammar Approach*. PhD thesis, The Ohio State University, 1992.
- Mary Dalrymple. *Lexical Functional Grammar*, volume 34 of *Syntax and Semantics*. Academic Press, New York, 2001.
- Mary Dalrymple and Tracy Holloway King. Nested and crossed dependencies and the existence of traces. In Tracy Holloway King and Valeria de Paiva, editors, *From Quirky Case to Representing Space: Papers in Honor of Annie Zaenen*. CSLI Publications, Stanford, Ca, 2013.
- D. Flickinger and J. Nerbonne. Inheritance and complementation: a case study of *easy* adjectives and related nouns. *Computational Linguistics*, 18.3:269–310, 1992.
- Jacqueline Gueron and Robert May. Extraposition and logical form. *Linguistic Inquiry*, 15(1):1–32, 1984.
- Rodney Huddleston and Geoffrey K Pullum, editors. *The Cambridge grammar of the English language*. Cambridge University Press, Cambridge, 2002.
- Paul Kay and Ivan A. Sag. Cleaning up the Big Mess: Discontinuous dependencies and complex determiners. In Hans C. Boas and Ivan A. Sag, editors, *Sign-Based*

- Construction Grammar*, number 193 in CSLI Lecture Notes, pages 229–256. CSLI Publications, 2012.
- Jong-Bok Kim and Ivan A. Sag. English object extraposition: A constraint-based approach. In Stefan Müller, editor, *The Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar, Department of Informatics, University of Lisbon*, pages 192–212, Stanford, 2005. CSLI Publications.
- Jong-Bok Kim and Peter Sells. The Big Mess construction: Interactions between the lexicon and constructions. *English Language and Linguistics*, 15:335–362, 2011.
- Carl J. Pollard and Ivan A. Sag. *Head-Driven Phrase Structure Grammar*. University of Chicago Press, Chicago, 1994.
- William J. Poser. Blocking of phrasal constructions by lexical items. In Ivan Sag and Anna Szabolsci, editors, *Lexical Matters*, pages 111–30. CSLI, Stanford, 1992.
- Louisa Sadler and Doug Arnold. Prenominal adjectives and the phrasal/lexical distinction. *Journal of Linguistics*, 30:187–226, 1994.
- Ivan A. Sag. Sign-Based Construction Grammar: An informal synopsis. In Hans C. Boas and Ivan A. Sag, editors, *Sign-Based Construction Grammar*, number 193 in CSLI Lecture Notes, pages 69–202. CSLI Publications, 2012.
- Ida Toivonen. *Non-Projecting Words. A Case Study of Swedish Particles*. Kluwer, Dordrecht, 2003.
- Frank Van Eynde. An HPSG treatment of *it*-extraposition without lexical rules. In G. Durieux, W. Daelemans, and S. Gillis, editors, *CLIN VI. Papers from the Sixth CLIN Meeting*, pages 231–248. University of Antwerp, Center for Dutch Language and Speech, 1996.
- Frank Van Eynde. Prenominals in Dutch. In Jong-Bok Kim and Stephen Wechsler, editors, *The Proceedings of the 9th International Conference on HPSG*, Stanford University, 2003. CSLI Publications.
- Frank van Eynde. The Big Mess construction. In Stefan Müller, editor, *The Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*, pages 415–433, Stanford, 2007. CSLI Publications.