Arbeitspapiere des SFB 340
Bericht Nr. 132, September 1999

SOME THOUGHTS ON TOUGH MOVEMENT

MIKE CALCAGNO

Contents

Adam Przepiórkowski suggested substantial revisions to an earlier version of the paper. Undoubtedly, the paper is much improved because of this. I owe a special debt to Tom Hukari and Bob Levine, who have greatly influenced my thinking on, and appreciation for, the subject of this paper. Thanks also to Ivan Sag, for allowing me to take a look at some of his preliminary work on related subjects, and to Gerald Penn, for help with some of the data and argumentation. Helpful comments on earlier drafts were also given by Frank Richter, Detmar Meurers, Christian Richard, and the members of the Winter 1999 Seminar on missing object constructions at Ohio State, which included Martin Jansche, Bob Levine, Carl Pollard and Nathan Vailette.

ISSN 0947-6954/99

1. Introduction

Tough movement, originally formulated by Rosenbaum 1967, refers to an early transformational rule in which a certain class of predicates (mostly adjectives like 'easy', 'hard', 'impossible') licenses an operation by which a sentence like (1a) is derived from one like (1b).

(1) a. Kim was easy/hard/impossible for me to please.
    b. It was easy/hard/impossible for me to please Kim.

In non-transformational theories like HPSG (Pollard & Sag 1994), one can take the term movement as a metaphor for some connectivity mechanism which allows the material in the subject of the (a)-sentence above, which I will refer to as the tough subject, to satisfy the selectional or semantic requirements of some other element in the utterance, at a distance which appears to be unbounded.

(2) Kim was hard for me to persuade Robin to try to deal with –.

More generally, most recent theories of unbounded dependency constructions (UDCs) concern themselves with associating overtly-realized elements ("fillers"), elements that have "moved" away from the place where they "should" have been realized, with phonologically empty place-holders ("gaps" or "traces") that actually occur locally to the element that is imposing the relevant selectional requirements. Whether or not gaps actually exist, in the sense of being psychologically real or linguistically necessary, is a matter of considerable debate (see, e.g., Fodor and Sag, 1994), and I am going to take the position in this paper that they do not. I will, however, use the term "gap" metaphorically, and indicate places where tough-subjects "come from" by drawing a hyphen (–) in places where I think doing so contributes to clarity.

Whether tough movement (which I will sometimes just call "tough") should or should not be treated as a UDC has also been a controversial matter in the literature to date. In early transformational work, notably by Postal, tough was not treated as a UDC at all, but rather as a species of complement object deletion. This view survived for a number of years, although it was not uncontroversially accepted. Beginning with Chomsky 1977 and in subsequent GB literature, attempts were made to assimilate tough movement to Wh-movement, and, although at some point the distinction was made between movements where the filler occurs in a non-argument ($\overline{\text{A}}$) versus argument (A) position, this view of tough goes relatively unchallenged at least through Chomsky 1986. Cinque 1990, however, does not treat tough as a UDC at all, and instead claims that tough gaps are best treated as instances of an empty pronoun, *pro*.

In more recent non-transformational approaches, beginning with Gazdar et al. 1985, the view that tough is a kind of UDC is certainly in the majority. In HPSG, Pollard & Sag 1994 treat missing object constructions as a kind of unbounded dependency, and a recent proposal by Bouma et al. 1998 pretty much duplicates that proposal. Grover 1995, however, has recently argued that tough constructions are "really very unlike ordinary UDCs" and "have much more in common with control constructions."

What is at issue are some similarities and differences between tough movement and certain other UDCs (which seem more well-behaved), and how best to handle this mixed bag of properties. The two theoretical choices, it seems, will always be (a) to treat it as some kind of UDC and try to explain the differences (e.g., for most speakers, tough gaps cannot occur as the subject of an embedded finite clause), or (b) to treat it as something else and try to explain the similarities (e.g., tough movement is apparently unbounded; it licenses parasitic gaps, gives rise to strong crossover effects, and exhibits sensitivity to island constraints).

I do believe that there are reasons to choose one alternative over the other, and the defense of tough movement as a UDC is the main goal of this paper. Of course, I have my own take on how then to explain the differences between tough and other kinds of UDCs, and I present an account in terms of the traceless theory of extraction proposed by Bouma et al. 1998, and show how my account can handle some of the problematic cases raised by Cinque and Grover.

2. Defending the UDC Position

As alluded to above, there has been some controversy recently as to whether tough movement is best treated as an unbounded dependency.

There are, as some authors have pointed out, certain idiosyncratic properties of these constructions which at least suggest that the connection between tough and other kinds of UDCs (e.g., Wh-movement, topicalization) might be weak enough to justify a completely different means by which tough subjects are associated with their corresponding gaps.

I would like to contend, however, that there are certain properties of tough which suggest, at least, that the null hypothesis should be that tough and things like Wh-movement are species of the same thing.

Parasitic gap licensing: Tough movement, like Wh-movement, licenses parasitic gaps.

(3) a. These color-coded reports should be easy to file – without having to read –.
b. Which reports did you file without having to read?

(4) a. Kim is easy to believe friends of – admire –.
b. Who do friends of admire?

Parasitic gap licensing is often taken as a diagnostic for whether or not an element actually "moved." Note that neither passive nor raising, nor some combination of the two, give rise to the appropriate effect.

(5) a. *That report was filed by Sandy without having to read.
b. *Kim believed that report to be filed without having to read.
c. *That report seemed to be filed without having to read.

The most reasonable conclusion to draw is that whatever it is about Wh-movement that licenses parasitic gaps, also licenses parasitic gaps in the case of tough movement.

Island sensitivity: Tough movement is impossible out of constituents which are widely recognized to be islands for extraction.

(6) a. *Kim would be difficult to imagine a person who dislikes –. (from Hukari & Levine 1987)
b. *Who did Kim see a person who dislikes –?

(7) a. *Kim would be difficult to read the paper without Bill consulting –.
b. *Who did Kim read the paper without Bill consulting –?

(8) a. *Kim is tough to like Sandy and –.
b. *Who did Kim see Sandy and –?

There are, of course, well-known exceptions to some of these constraints. Strikingly, tough patterns closely with Wh-movement with respect to the counterexamples, as well.

(9) a. Kim is hard to take a picture of –.

b. Who did you take a picture of –.

(10) a. That book might be hard to go to London without reading –.

b. Which book did you go to London without reading –?

Other kinds of "movement", e.g., passive and raising, fare considerably worse in these cases.[1]

(11) a. ?*I was taken a picture of – by tourists yesterday.

b. *That book was travelled to London without reading –.

c. *I believe John$_i$ a picture of $-_i$ to be on display in the museum.

d. *John seems a picture of – to be hanging in the museum.

Again, the most reasonable conclusion to draw, it seems, is that whatever is ruling out the ill-formed Wh-movement in (6) is also responsible for ruling out tough movement from the same type of constituent. This would be surprising if Wh- and tough movement were radically different operations.

These two properties constitute what I take to be the most compelling evidence in favor of treating tough as a UDC. Any account which does not must come up with a convincing alternative explanation for the effects just described.

In the remainder of this section, I review two such proposals. The first, Cinque 1990, argues that the gap in tough constructions, as well as parasitic gaps and the gaps in legal adverbial island extractions, are actually instances of the empty pronominal *pro*, and thus different from gaps associated with ordinary UDCs. The second, Grover 1995, proposes that tough constructions, along with other missing object constructions, are really a kind of control construction, and thus not related to UDCs at all.[2]

I will try to show that Cinque and Grover do not provide any truly compelling evidence for rejecting a UDC analysis of tough, and that their own accounts of the facts above suffer from enough empirical and conceptual

[1] These are not meant to be shocking pieces of data.

[2] Grover's account is intended to apply not only to tough constructions, but also to purpose infinitives ('He brought John along to talk to'), *too/enough* complements ('That book is too long to read in one go') and *need* predicates ('The car needs washing'). I am not sure that these constructions form a natural class (the *need* predicates seem altogether different, for example), and, for the purposes of this paper, I will focus only on tough.

shortcomings to warrant rejecting them. Of course, the points these authors raise regarding apparent differences between tough and more standard kinds of UDCs are well-taken. It is my position, however, that these differences are not incompatible with a UDC analysis, and, in light of the striking similarities between tough and UDCs outlined above, I propose that the most reasonable conclusion is that tough is a kind of UDC.

2.1. **Cinque 1990.** Cinque's argument for rejecting a UDC analysis of tough is actually twofold, consisting first of the claim that missing object constructions form a single class with adjunct-island extractions and parasitic gaps, as attested by their common distributional idiosyncracies, and second that the idiosyncracies associated with these three phenomena are all explainable if one takes the missing element in these constructions to be an instance of the empty pronominal *pro* (cf. Suñer 1982) rather than a Wh-trace.[3] That is, what Cinque presents is actually (a) a set of empirical claims about the properties of these constructions, properties which, he argues, in no way follow from a treatment of the constructions as a UDC; and (b) an explanation of these putative facts in terms of *pro*.

Here are the empirical claims:

- *(i):* Missing object constructions, extraction from adjunct-islands, and parasitic gaps are all restricted to instances with NP fillers.
- *(ii):* All three constructions are barred from associating a filler with a gap in subject position of a finite clause;
- *(iii):* The second object of the dative construction is inaccessible to extraction in the family of constructions under consideration, though evidently completely accessible to normal wh-extraction;
- *(iv):* The class of extractions under discussion cannot extract an NP from an adjunct that itself occurs within an adjunct phrase.

I will argue against each of these one at a time, focussing primarily on how the argument pertains to tough.

2.1.1. *Claim (i).* The claim that tough-subjects must be NPs appears to be false. That is, the following examples are uncontroversially acceptable:

[3] Cinque's analysis of parasitic gaps was argued against at length in Calcagno, Hukari and Levine (CHL) pear, and many of the arguments presented there also apply to tough. Some of the arguments in this subsection, including many examples, are due to CHL 1998.

(12) a. That Robin is a spy would be reasonable to conjecture at this point.
b. For Robin to be here on time would be reasonable to expect.
c. Him getting here on time is hard to imagine.
d. Getting here on time is hard to manage when one doesn't have a car.
e. To make mistakes on the first problem set is reasonable for one to expect in a course this hard.
f. Proud of themselves is hard for me to picture ANY of them feeling, after what they've all done.
g. That such a tradition lies behind The Iliad and Thy Odyssey, at least, is hard to deny –.[4]

Note that, in almost all of the cases above, the verbs whose objects are missing are not readily interpretable (in the relevant sense suggested by the examples) in the absence of the gapped material; that is, it is doubtful that these examples can be explained away as not involving extraction. And, particularly in (12e,f), there seems to be no plausible way to interpret the extracted material as an NP. In some of the cases above, it is at least plausible that some sort of accommodation may allow the clauses in question to precede a pronoun. That is, the following continuations of the (a-e) examples are either acceptable or close to acceptable:

(13) a. That Robin is a spy would be reasonable to conjecture at this point, and thus I conjectured ??it.
b. For Robin to be here on time would be reasonable to expect, and I expected it.
c. Him getting here on time is hard to imagine, but I imagined it anyway.
d. Getting here on time is hard to manage when one doesn't haven't a car, but Robin always seems to manage it.
e. To make mistakes on the first problem set is reasonable for one to expect in a course this hard, so I learned to expect it.

However, the (f) example, and others like it, fails quite miserably in parallel examples.

(14) a. *Proud of themselves is hard for me to picture ANY of them feeling, after what they've all done, but they felt it anyway.

[4]I am grateful to Martin Jansche for this example. He found it in the Brown corpus.

The situation, of course, is not without complexity. That is, the following are quite unacceptable:

(15) a. *Under the table is reasonable to keep your beer.
b. *To John is easy to give a book.
c. *Carefully is easy to proceed.
d. *There is easy to expect to be a unicorn in the garden.
e. *It is easy to expect to rain./It is easy to expect it to rain.

It is unclear, however, whether the badness of these examples has anything to do with tough movement. The subjects in these cases sound odd with a predicative adjective of any sort, as the examples in (16) illustrate.

(16) a. *Under the table is ablaze/reasonable/surprising/wonderful.
b. *To John is surprising/wonderful.
c. *Carefully is unacceptable/atrocious/reasonable.
d. *There is surprising.
e. *It (expl) is surprising.

Contrast those examples with the data in (17).

(17) a. That Robin is a spy was surprising.
b. For Robin to be here on time would be surprising/wonderful.
c. Him getting here on time is surprising.
d. Getting here on time is wonderful.
e. To make mistakes on the first problem set is surprising.

Note that some of the ill-formed examples in (16) can be improved if they are predicated of NPs, as in the following examples:

(18) Under the table is the safest place in the house.

And the examples in (19), which involve tough-adjectives, albeit inside an NP, seem perfectly acceptable.

(19) a. Under the table is hard to imagine being a good place to store beer.
b. Carefully would be the easiest way to proceed at this point.

Finally, we should note that the restriction of the tough gap to an NP, even if it were true, does not necessitate an analysis of tough as a non-UDC. Restrictions of this sort are not unique to tough (or to the other two classes of data Cinque considers). For example, it is well-known that, while NP objects can be extracted out of locative adjuncts to nominal heads, extraction of the whole PP is forbidden, as the example in (20) illustrates.

(20) a. I'm looking forward to seeing the bridge over the Neckar.
 b. ?Which river are you looking forward to seeing the bridge over?
 c. *Over which river are you looking forward to seeing the bridge?

Only a UDC analysis of these examples seems plausible. That is, Cinque's non-UDC analysis of islands in a tough construction, which relies on LF pied-piping of the constituent containing the tough gap, is infeasible as an account of these examples. LF pied-piping of the PP in (20c) is extremely dubious, given the impossibility of S-structure pied-piping, while pied-piping of the NP, assuming the expanded version of g-projection Cinque introduces to accomodate relative clauses, is ruled out by the fact that NPs cannot pied-pipe in interrogative clauses:

(21) a. ?I wonder of whom Robin took a picture.
 b. *I wonder a picture of whom Robin took.

The fact that none of the g-projections of the gap site in (20) is interrogatively extractable in the overt syntax of English (cf., '*To visit the bridge over which river are you looking forward to seeing?') makes it necessary to posit some source other than the nature of the gap for the NP-exclusiveness of extraction from such contexts. It is not unreasonable, then, to assume that this same constraint might be brought to bear on tough constructions.

2.1.2. *Claim (ii).* Cinque's explanation for why the gap in a tough construction cannot be a finite subject depends on certain GB-internal assumptions concerning the "identification" of pronouns. The argument can be roughly summarized as follows (adapted directly from CHL):

> If *pro* can only appear in syntactic contexts which allow it to be uniquely identified with respect to its ϕ features, then since it cannot remain in finite subject position (since neither I nor C need be taken to be 'proper licensors' in Cinque's terms, just as they are not proper head governors for purposes of the ECP) and since, if it moves to Spec of CP, it will no longer be in an A-position and hence cannot be interpreted at S-structure as a variable, there is really nowhere it can be as a finite subject that will give rise to a legal structure.

First, there is some doubt as to whether the empirical basis for the claim is correct, given examples like those in (22), which I find close to acceptable.

(22)
  a. ??Who is easy to believe likes ice cream?

And this restriction is certainly untrue of the other two types of extraction that, for Cinque, form a natural class with tough. Consider the examples in (23).

(23) a. Which people did you invite to the party without thinking would actually come?
  b. Which guru did Robin travel halfway around the world to PROVE had been cheating young students out of their money?
  c. Robin is someone who even friends of believe should be closely watched.
  d. He was forced to deal with executives who even friends of know are backstabbers.

I am willing, however, to assume that tough-gaps are, in most cases, prohibited from the subject of finite clauses, and this appears to be a genuine difference between tough and UDCs like Wh-movement.

I do not, however, think that it is necessarily incompatible with a UDC analysis, and, in fact, it has been accounted for in the past, most notably in the UDC analyses of Gazdar, Klein Pullum & Sag 1985 and Pollard & Sag 1994, by simply assuming that NP gaps in this construction must be accusative.

2.1.3. *Claim (iii).* Cinque's third claim, that tough movement is impossible from second dative object position, appears, like his first claim, to be false.

According to Cinque, the prohibition on gaps in the second dative object position follows immediately if weak pronominals, such as 'it', and therefore *pro*, cannot occupy the second dative position, and he cites examples such as (24) to show this to be a true generalization:

(24) *I gave that man it/them.

Hence, on Cinque's account, the second dative object phenomenon receives a unified account on the assumption that *pro*, not Wh-trace is involved in the three constructions where extraction of this object is putatively forbidden.

This argument, unfortunately, turns out to be seriously flawed. First, Cinque does not actually consider the source of this phenomena in any detail, and he simply takes it as a given that the putative parellelism among the three constructions in question, as opposed to Wh-movement, further reinforces their classification as a natural class, and that the identification of *pro* as a kind of pronoun, prohibited in overt cases from appearing as the second object and therefore plausibly banned in its inaudible form, makes the *pro* solution optimal in this case. But in the most important consideration of blocked unstressed pronouns, that in Zwicky 1986, which Cinque was apparently unaware of, it is shown quite convincingly that the conditions determining where an unstressed pronoun may appear are ultimately reducible to strictly prosodic factors – in particular, to the requirement that no unaccented pronoun may form a prosodic phrase on its own.

Even ignoring Zwicky's alternative, it appears that the data in question fail to support Cinque's claim. Consider the following, where the extraction in question is quite acceptable, but the admissibility of an unstressed pronoun in the same position is bad.[5]

(25) a. I can think of a couple of presents that would be really fun to give Robin for his birthday.
b. *We gave Robin them for his birthday.

(26) a. Junk bonds are nearly impossible to sell ANYONE nowadays.
b. *We sold someone them.

(27) a. There's certain evidence that we may have to go to trial without ever showing Robin.
b. *We showed Robin it.

(28) a. In fact, there was certain crucial news about the case that Robin went on vacation without even telling her attorneys.
b. *Robin told her attorneys it.

Such examples seem to be entirely grammatical. The only conclusion, it appears, is that the content of Cinque's claim is simply not compatible with the facts of English.

[5]For the following judgements to make sense, assume that all the pronouns in the relevant examples are unstressed. One way to accomplish this while reading is to use a contracted form (e.g., 'em for them, and so forth).

2.1.4. *Claim (iv).* The prime piece of evidence here, as noted earlier, is the claim that parasitic and adjunct-island gaps, in common with tough constructions, "seem able to violate one, but not more than one, island at a time" (Cinque 1990:133). Cinque would argue that, since a Wh-movement analysis would supposedly lead one to expect that an arbitrary number of island-violating extractions should be possible if a single one is, then this fact is surprising if tough is an ordinary UDC. As alluded to somewhat obliquely in the discussion of Cinque's first claim, the analysis that Cinque suggests explains island violations in cases such as these not in terms of simple Wh-extraction, but rather a form of pied-piping in which Wh-properties attributed (without any readily apparent independent motivation) to *pro* are subject to a set of conditions on based on Kayne's 1983 connectedness proposal. It is the interaction between the LF pied-piping and connectedness that Cinque claims gives rise to the observed one-island-only effect.[6]

There are reasons, however, to reject this argument entirely. It has been proposed numerous times in the literature that island effects display a cumulative pattern and that it is precisely the iteration of such violatons which causes difficulties that result in a negative assessment of the data. That is, Cinque's rejection of a conventional filler/trace link in the cases in question is unnecessary, given the existence of clear examples in uncontroversially *wh*-extractions which behave the same way. As far as I am aware, the first observation in print that simple iteration can result in an island effect appeared in Fodor 1983, where the following example is given:

(29) a. Who did you take a photo of?
b. *Who did you take a photo of a statue of?

Not long after Fodor's article appeared, Chomsky 1986 suggested that subjacency should not be treated as an all-or-nothing effect, but rather that cumulative occurrences of 1-subjacency violations might be thought of as additive, along lines very reminiscent of Fodor's speculations. More recently, Kluender 1998 has proposed a model of island effects as the accumulation of processing burdens, which, while still more qualitative than quantitative, presents a psycholinguistic architecture that makes cases like Fodor's (29), which Cinque finds so significant, an entirely straightforward matter calling for no special dichotomies in the classification of

[6] Unfortunately, space limitations prevent me from presenting the specifics of Cinque's proposal in any detail. The reader is referred to his book.

UDCs. That is, what Kluender's results show is that the data Cinque brings to bear on this question is completely compatible with what is known about the cumulative effect of processing difficulties, making the elaborate pied-piping scenario Cinque posits unnecessary.

In addition, there is excellent reason to reject Cinque's LF pied-piping hypothesis on strictly empirical grounds. The basic data here are cases of perfectly well-formed missing object, parasitic and adjunct island gaps that do not correspond to legal pied-pipable constituents in English.

(30) a. Which papers did you take before Robin discovered?
b. I wonder which papers Leslie escaped to Bermuda with without Robin ever discovering.
c. Those papers would be hard to imagine one leaving the country without Robin discovering.

Given the necessity of LF pied-piping that Cinque posits, one is apparently required to move the entire adjunct phrase, which is the maximal g-projection of the [+*wh*] feature originating in *pro*. But for many, and probably all, English speakers, such pied-piping is impossible:

(31) a. *The papers$_i$ before Robin discovered which$_i$ you took – are on the table.
b. *These are the papers Robin discovering which we must prevent.

In view of this widely acknowledged distributional fact about English (which contrasts very conspicuously with that of possessive gerundive phrases as in 'These are the papers Robin's discovering which we must prevent'), it is hard to see how Cinque's analysis can meet the plausibility requirement which he himself alludes to – namely, that claims of LF pied-piping need to be ratified by evidence that such pied-piping is possible in 'overt' syntax.

One thus finds, on closer examination of Cinque's arguments, that the specific factual claims which are used to support the various arguments given for *pro* rather than Wh-extraction in the constructions in question are simply mistaken.

Hence, on these grounds alone, it seems unecessary, and indeed incorrect, to abandon the default position that tough gaps (inter alia) are instances of the same mechanism as other filler/gap linkages.

2.2. **Grover 1995.** Grover's arguments for rejecting a UDC analysis of tough are in some senses very similar to Cinque's, and, indeed, many of the arguments against Cinque's proposal also apply to Grover. The basic form of her argument is that tough movement exhibits a cluster of properties which do not follow immediately from a UDC approach, and that these properties are better accounted for as treating tough movement as a kind of control construction. Those properties which Grover takes to be problematic for the UDC approach are as follows:

- *(i):* Tough gaps are prohibited from the subject position of a finite clause.
- *(ii):* The antecedent occurs in A position, not ($\overline{A}$) position.
- *(iii):* Tough movement is not truly unbounded, in that a tough gap cannot appear inside a finite clause.
- *(iv):* There is a non-parallelism between violable islands in UDCs and in tough movement.
- *(v):* There is an asymmetry between the acceptability of sentential "fillers" in UDCs on the one hand and in tough movement on the other.

I will skip (i) here, as this was already discussed in the previous subsection. In essence, I am willing to grant the facts on this one. I don't think, however, that these facts are totally incompatible with a UDC analysis of tough. It should be noted that what Grover actually claims on this point is much stronger than what I present in (i). That is, Grover (p. 84) actually states that "the 'gap' must always be an object gap," and the mechanics of her account suggests that she understands this in the rather narrow sense of being on the complements list of a verb or some argument that is selected by the verb (like a preposition). This does not seem right, since the complements of adjuncts, which are generally taken **not** to appear on valence lists, can undergo tough movement quite easily, as the examples below illustrate.

(32) a. That store is easy to walk to –.
b. Mondays are hard to wake up on –.
c. That book might be hard to go to London without reading –.

I will also pretty much skip over (ii), that the filler of a tough construction is in A, as opposed to $\overline{A}$ position. This seems to me to have no bearing on the UDC-hood of the tough construction. That is, clefts behave uncontroversially as UDCs, and one could argue that the "moved" constituent in this construction ends up in an argument position as well.

(33) a. What I believe Mary wanted was a deer.
b. It was under the table that John believed Mary would look first.

So what is the "real" filler in these cases? If it is *what* and *that*, then one could argue that the filler was not in an A position. But if it is *under the table* and *a deer*, then this argument does not hold. A similar situation arises with tough. If we take the tough subject as the real filler of the tough gap, then it is in an A position. However, I will argue, following Pollard & Sag 1994, that the tough subject is only semantically associated with the tough gap, and that the actual filler is supplied by the tough predicate. In this way, it is treated quite similarly to clefts (which, again, are uncontroversially UDCs).

Grover's argument in (iii) appears to be empirically incorrect. That is, the following examples, to my ear, are impeccable.

(34) a. That problem is easy to think you understand.
b. Chess is probably harder to believe anyone has mastered.
c. Monica Lewinsky might be fun to say one saw.
d. This kind of mistake is hard to realize that you're making –.

It should be noted that the mechanics of Grover's analysis make the data above problematic. That is, on Grover's account, the gapped downstairs verb would have two subjects, one of which must be realized to form the sentential complement of the first verb under the tough predicate. If this is the case, though, then Grover's revised head-subject schema (p. 93), which requires that the SUBJ value of the head daughter of a *head-subj-struc* be a one-member list, must be dropped. But if this is the case, then there is nothing preventing a verb from combining with two subjects.

Grover's argument (iv) appears also to be empirically incorrect. In the previous subsection, I cited the following examples which I took to make the opposite point.

(35) a. Kim is hard to take a picture of –.
b. Who did you take a picture of –.

(36) a. That book might be hard to go to London without reading –.
b. Which book did you go to London without reading –?

Grover's data on this point appears below (the judgements are hers).

(37) a. Who did you sell some pictures of –?

b. Which country did George meet the king of –?

(38) a. Sandy was easy to take pictures of –.
b. *Sandy is hard to sell some pictures of –.
c. *France would be impossible to meet the king of –.

Island violability is a fragile phenomena, and I suspect that what is going on here, if you accept Grover's judgements, has more to do with the sensitive semantic and processing conditions involved with island violations than with the UDC-hood of tough movement. Note that the ill-formed examples above are easily improved (note that I would have marked Grover's (37b) with a ?, as well).

(39) a. Sandy was the hardest to sell some pictures of –.
b. ?A country that large is hard to meet the king of –.

Finally, Grover's argument (v), that sentential complements in a UDC behave differently than tough subjects, is also slightly off the mark. The following utterances seem perfect to me, especially with constrastive focus on the matrix subject (indicated clumsily here by capital letters).

(40) a. That I solved the problem, NO ONE could believe.
b. Whether John solved the problem himself, no one would ever know.

The oddness of some of these examples certainly has more to do with the semantics of preposing than with the UDC-hood of the tough construction. Note that moving a sentential complement in a cleft construction is usually quite good.

(41) a. It's that he didn't want to do it my way that bothered me the most.
b. What I thought was that he would play the drums.

It is hard to claim that a UDC is not involved in these examples, and thus the relevance of Grover's claim, even if empirically correct for some cases of tough movement, to the argument that tough is not a UDC is unclear.

3. A UDC analysis of Tough movement

In this section, I provide a UDC analysis of tough movement within the framework of HPSG. It improves on the analyses of Bouma et al. 1998 and Pollard & Sag 1994, and addresses some of the differences between tough and ordinary Wh-movement that have emerged from our discussion of Cinque 1990 and Grover 1995. Thus, the account here can be taken to show (as was argued previously) that these kinds of facts are not incompatible with a UDC approach.

The theory of UDCs on which my analysis is based is essentially the traceless analysis given in Bouma et al. 1998, although I find it necessary to spell out in more detail some aspects of the analysis that those authors chose not to focus on, and, on the other hand, to gloss over other parts which are not immediately relevant to the task at hand. Unfortunately, space limitations force me to assume some familiarty with the analysis of extraction in Pollard & Sag 1994, and in subsequent HPSG literature. I will try, of course, to refer readers to the appropriate places there and in Bouma et al.'s manuscript in cases where I think the material is particularly relevant.

3.1. **Without a trace.** Although a great majority of contemporary theories of UDCs posit the existence of traces, there are, as Bouma et al. points out, quite a few traceless approaches that have been proposed in recent years. Theories along these lines have cropped up in a diverse body of work, including phrase-structural approaches like GPSG (Gazdar et al. 1985); deductive formalisms like flexible categorial grammar (Ades & Steedman 1982, Steedman 1996) and type-logical grammar (Morrill 1994); and "unification-based" approaches like LFG (Kaplan & Zaenen 1989) or HPSG (Pollard & Sag 1994, chapter 9).[7]

An appealing feature of this body of work is that it seems to accord with recent arguments by Sag 1998 and Sag & Fodor 1994, which hold that there is little or no independent evidence as to the existence of these invisible constituents. In brief, Sag 1998 points out that trace-based accounts of phenomena such as *wanna*-contraction, auxiliary contraction, quantifier floating, as well as weak and strong crossover, which either are, or have been at one time, taken to argue in favor of the existence of trace, have either been shown to be seriously flawed, or have been superseded by superior analyses which do not appeal to traces at all. Furthermore, Sag 1998 argues that there exists psycholinguistic evidence (see Pickering

[7]As far as I am aware, there is no such approach in the GB/minimalist literature (in which, if anything, the number of empty elements – and their role in the theories which are currently in vogue – is growing at a rapid rate).

& Barry 1991, Sag & Fodor 1994) which suggests that such elements cannot be psychologically real, as well as additional linguistic evidence which to date has either resisted or is seemingly incompatible with trace-based accounts.

3.2. **A traceless theory of UDCs.** Bouma et al.'s analysis of UDCs is distinguished not only by its lack of dependence on trace, but also by significant and important differences with other, earlier, traceless analyses.

First, it is distinguished from the categorial accounts by its use of a feature to locally encode information about the filler all along the extraction path. This allows an account for a range of cross-linguistic phenomena which appears to be manifest only along such paths (e.g., as attested in Irish (McCloskey 1979), French (Kayne & Pollock 1978), Icelandic (Maling & Zaenen 1978), *inter alia*).

Second, it is distinguished from previous G/HPSG accounts by two key features: (a) it provides a uniform account of extraction of subject, complement and adjunct, a uniformity which has been argued for on an empirical basis at length by Hukari & Levine 1995, 1996; and (b) it does not involve the use of lexical rules or meta-rules, thus making it conceptually more appealing to many HPSG practicioners for whom lexical rules, especially of the type proposed to handle extraction, represent an unwanted complication of HPSG's underlying formalism.

3.2.1. *Complement extraction.* Taking the feature geometry of Pollard & Sag 1994 as a starting point, the following modifications must be made in order for the theory presented here to work.[8]

(i): The valence feature SUBCAT, whose value was in Pollard & Sag a list of *synsem* objects, is replaced by the features VAL(ENCE), ARG(UMENT)-ST(RUCTURE), and DEP(ENDENT)S. The value of the first feature is a *val(ence)-obj(ects)*, and the value of the latter two are lists of *synsem* objects. Bouma et al. is responsible for the introduction of DEPS, and they restrict it to words. The sort *val-obj* has the features SUBJ, SPR and COMPS, whose values are lists of *synsem*.

[8]Some of these changes have been argued for by numerous authors over the course of the last few years, some were specifically proposed by Bouma et al.; I tweak things slightly here.

In the new setup, VAL encodes the combinatoric potential of words and phrases, and it is this feature that is constrained by the analog of Pollard & Sag's subcategorization principle; ARG-ST is, in my conception, something like an LFG f-structure, and it plays a crucial role in the binding theory; DEPS is central to Bouma et al.'s theory of UDCs, as well as to the theory of case presented in Przepiórkowski 1996, which is compatible with the approach I present below.

(ii): The feature BIND, which is appropriate for *category* objects and takes a set of *local* objects as value, replaces the nonlocal feature TO-BIND, following Pollard & Yoo 1998 and Manning & Sag 1995. It plays a role in the lexical binding of the nonlocal feature SLASH. Its value will be the empty set for all words except tough predicates.

In a theory of UDCs, the gap is subject to certain local restrictions (e.g., the verb 'put' requires that one of its objects be a PP), and it occurs in a place in the structure where such restrictions are standardly assumed to apply.[9] The filler, in turn, is also subject to these restrictions, even though it occurs in a place in the structure which is at a distance from where the restriction would be usually taken to apply.

Feature-based accounts traditionally accomplish the above by "percolating" information, via a feature like SLASH, from the gap site (usually a daughter realized as a trace) to the place where the filler is actually realized in the tree, subject to principles governing this percolation (e.g., the Foot Feature Principle (FFP) of GPSG or the Nonlocal Feature Principle (NLFP) of HPSG). Both the FFP and the NLFP are configurational in the sense that they constrain SLASH percolation in terms of mothers and daughters in a phrase structure tree.

Bouma et al. suggest an alternative, in which the analog of NLFP is replaced by two constraints: (a) a lexical constraint on "SLASH amalgamation", and (b) a constraint on phrases which governs the "inheritance" of SLASH between a particular head and its mother. These constraints are given in (42) and (43).

[9] In a traceful analysis, the gap appears in a local tree with its selector. In traceless analyses, it appears on (one of) the selector's valence lists.

(42) SLASH Amalgamation Constraint (SLAC) (Bouma et al. 1998)

$$word \rightarrow \left[\text{SS}\left[\begin{array}{ll}\text{LOC} & \left[\text{CAT}\left[\begin{array}{ll}\text{DEPS} & \left\langle[\text{SLASH }\boxed{1}], \ldots, [\text{SLASH }]\right\rangle \\ \text{BIND} & \boxed{0}\end{array}\right]\right] \\ \text{NONLOC} & \left[\text{SLASH }(\boxed{1} \cup \ldots \cup\ ) - \boxed{0}\right]\end{array}\right]\right]$$

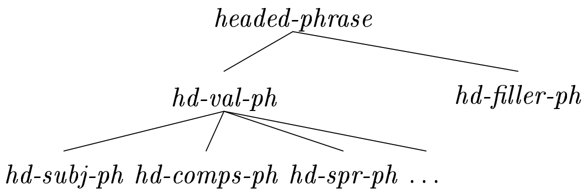
(43) SLASH Inheritance Principle (SLIP) (Bouma et al. 1998)

$$hd\text{-}val\text{-}ph \rightarrow \left[\begin{array}{l}\text{DTRS HEAD-DTR SS NONLOC SLASH }\boxed{1} \\ \text{SS NONLOC SLASH }\boxed{1}\end{array}\right]$$

SLASH amalgamation ensures that, if a dependent on some element is slashed, the head which selects it will also be slashed. SLASH inheritance simply ensures that slash information is shared between the head daughter and its mother. The idea, looking ahead a bit, is that a gap will originate on some head's dependents list, and introduce a slash value. SLAC will ensure the head itself will have an identical, and therefore nonempty, SLASH value, and SLIP ensures that this value will be passed up to the head's mother, and so on up the head path of the structure.

Note that SLIP is a constraint on *hd-val-phrases*, which Bouma et al. take to be phrases involving head, complement, or subject daughters (and presumably specifier and adjunct daughters) but, crucially, no filler daughters. The division of headed phrases into *head-val-ph* and *hd-filler-ph* is new. The following piece of sort hierarchy makes the change explicit:

(44)



Objects of sort *hd-filler-ph* are subject to a different constraint, namely the one in (45):

(45) $$hd\text{-}filler\text{-}ph \rightarrow \left[\begin{array}{ll}\text{SS NONLOC SLASH }\boxed{2} \cup \boxed{3} & \\ \text{DTRS} & \left[\begin{array}{ll}\text{HEAD-DTR SS NONLOC SLASH }\boxed{2} \cup \left\{\boxed{1}\right\} & \\ \text{FILLER-DTR SS} & \left[\begin{array}{l}\text{LOC }\boxed{1} \\ \text{NONLOC SLASH }\boxed{3}\end{array}\right]\end{array}\right]\end{array}\right]$$

Here, the SLASH value of a given head/filler phrase is constrained to be the same as the SLASH value of the head daughter, minus the local value of the filler (plus any slash the filler might have had).[10] Intuitively, the idea here is that an element in SLASH gets removed (bound off) when an appropriate filler is encountered.

So, we now know how a SLASH value (corresponding to a gap) on a word's dependents list gets associated with and bound off by a filler. But how does a SLASH get on such a list in the first place? In previous HPSG analyses involving trace, a trace realized as a daughter would introduce the slash, and it would be passed up via the NLFP to the phrase which immediately dominated it. In previous traceless analyses, the gapped element was somehow removed from the verb's valence lists by way of lexical rule with its local value going into that verb's SLASH.

In Bouma et al., neither traces nor lexical rules are used. Instead, a constraint which governs the mapping between an element's dependents and the valence features SUBJ and COMPS is posited. This constraint, given in (46), allows some dependents to be realized as real valents and some to be realized as gaps.

(46) Dependent Realization (Bouma et al. 1998)

$$\textit{word} \rightarrow \left[\text{SS LOC CAT} \left[\begin{array}{ll} \text{VAL} & \left[\begin{array}{l}\text{SUBJ } \boxed{1} \\ \text{COMPS } \boxed{2} \ominus \textit{list(gap)}\end{array}\right] \\ \text{DEPS} & \boxed{1} \oplus \boxed{2}\end{array}\right]\right]$$

Dependent Realization requires that the first element on an element's dependents list be realized as the subject, with the rest of the dependents being realized either on the COMPS list or as a gap. Gaps are distinguished from overt elements along the lines of Sag 1997. That is, two subsorts of *synsem* are introduced, one for gaps and one for "canonical" elements. The following piece of sort hierarchy makes this explicit.

(47)
```
       synsem
      /      \
canon-ss   gap-ss
```

[10]Since extraction from fillers is in general excluded in English, as in (i), we may want to require that the SLASH value of the filler be empty.

(i): The boy from New York, I liked./*From NY, the boy, I liked.

We do have to ensure, however, that, in the case of parasitic gaps, extraction from "slashed" constituents is allowed, as in the utterance below.

(ii): This is the kind of misguided analysis that I never know what aspects of – to call attention to – first.

Gap-synsems introduce slashed elements onto SLASH by way of the following constraint:

(48) $$\textit{gap-ss} \rightarrow \begin{bmatrix} \textsc{loc}\ \boxed{1} \\ \textsc{nonloc slash}\ \left\{\boxed{1}\right\} \end{bmatrix}$$

Here, the *local* value of the gapped element is constrained to be identical to a singleton member of its SLASH value. If a gap-synsem object appears on, say, a verb's dependents list, this local value will be amalgamated to the slash value of the verb and passed up to the verb's mother, and then to the verb's mother's mother and so forth, until it is removed in a hd-filler-ph by the presence of a filler with an identical local value.

As an example, of how the system works, consider the partial description for the verb 'pleased' in (49), and the tree corresponding to the sentence 'Robin, I pleased' in (50).

(49) Partial description for the word object 'pleased', with gapped object NP

$$\begin{bmatrix} \textit{word} \\ \textsc{phon}\ \textit{pleased} \\ \textsc{ss} \begin{bmatrix} \textsc{loc cat} \begin{bmatrix} \textsc{val} \begin{bmatrix} \textsc{subj} \left\langle \boxed{3} \begin{bmatrix} \textit{canon-ss} \\ \textsc{loc}\ \text{NP} \\ \textsc{nonloc slash}\ \boxed{1} \end{bmatrix} \right\rangle \\ \textsc{comps}\ \langle\rangle \end{bmatrix} \\ \textsc{deps} \left\langle \boxed{3}, \begin{bmatrix} \textit{gap-ss} \\ \textsc{loc}\ \boxed{2} \\ \textsc{nonloc slash}\ \left\{\boxed{2}\right\} \end{bmatrix} \right\rangle \end{bmatrix} \\ \textsc{nonloc slash}\ \boxed{1} \cup \left\{\boxed{2}\text{NP}\right\} \end{bmatrix} \end{bmatrix}$$

Note that the distribution of its NP dependents is consistent with the Dependent Realization constraint in (46).

(50)

$$\begin{bmatrix} \textit{hd-filler-ph} \\ \text{PHON } \textit{Robin, I pleased} \\ \text{SS} \begin{bmatrix} \text{LOC CAT VAL} \begin{bmatrix} \text{SUBJ } \langle\rangle \\ \text{COMPS } \langle\rangle \end{bmatrix} \end{bmatrix} \\ \text{NONLOC SLASH } \{\} \end{bmatrix}$$

F

$$\begin{bmatrix} \text{PHON } \textit{Robin} \\ \text{SS LOC } \boxed{2} \end{bmatrix}$$

H

$$\begin{bmatrix} \textit{hd-sbj-ph} \\ \text{PHON } \textit{I pleased} \\ \text{SS} \begin{bmatrix} \text{LOC CAT VAL} \begin{bmatrix} \text{SUBJ } \langle\rangle \\ \text{COMPS } \langle\rangle \end{bmatrix} \end{bmatrix} \\ \text{NONLOC SLASH } \left\{\boxed{2}\text{NP}\right\} \end{bmatrix}$$

S

$$\begin{bmatrix} \text{PHON } \textit{I} \\ \text{SS } \boxed{3} \end{bmatrix}$$

H

$$\begin{bmatrix} \textit{word} \\ \text{PHON } \textit{pleased} \\ \text{SS} \begin{bmatrix} \text{LOC CAT} \begin{bmatrix} \text{VAL} \begin{bmatrix} \text{SUBJ} \left\langle \boxed{3} \begin{bmatrix} \textit{canon-ss} \\ \text{LOC NP} \\ \text{NONLOC SLASH } \langle\rangle \end{bmatrix} \right\rangle \\ \text{COMPS } \langle\rangle \end{bmatrix} \\ \text{DEPS} \left\langle \boxed{3}, \begin{bmatrix} \textit{gap-ss} \\ \text{LOC } \boxed{2} \\ \text{NONLOC SLASH } \left\{\boxed{2}\right\} \end{bmatrix} \right\rangle \end{bmatrix} \\ \text{NONLOC SLASH } \left\{\boxed{2}\text{NP}\right\} \end{bmatrix} \end{bmatrix}$$

3.2.2. *Subject and adjunct extraction.* Subject extraction requires no further modifications to the theory. The only difference is that, in accordance with Dependency Realization, a gap-synsem subject will be realized on the valence SUBJ list. The description in (51) illustrates.

(51) Partial description for 'please', with gapped subject NP

$$\begin{bmatrix} \textit{word} \\ \text{PHON } \textit{please} \\ \text{SS} \begin{bmatrix} \text{LOC CAT} \begin{bmatrix} \text{VAL} \begin{bmatrix} \text{SUBJ} \left\langle \boxed{3} \begin{bmatrix} \textit{gap-ss} \\ \text{LOC } \boxed{2} \end{bmatrix} \right\rangle \\ \text{COMPS} \left\langle \boxed{4}\text{NP}[\text{SLASH } \boxed{1}] \right\rangle \end{bmatrix} \\ \text{DEPS} \left\langle \boxed{3},\ \boxed{4} \right\rangle \end{bmatrix} \\ \text{NONLOC SLASH } \boxed{1} \cup \left\{\boxed{2}\text{NP}\right\} \end{bmatrix} \end{bmatrix}$$

The word object corresponding to this description licenses a VP with a singleton gap-synsem value as subject. Note that the local value of this gap will be amalgamated to the SLASH value of 'please' via SLAC (and passed up to any VP that 'please' is the head of via SLIP).

Of course, we have to ensure that this gap-synsem doesn't get realized as a real subject, and the following constraint does the trick:

(52) $\mathit{sign} \rightarrow \left[\textsc{ss}\ \mathit{canon\text{-}ss}\right]$

Verbs which are thought of as taking a sentential complement, like 'think', can instead select for a verbal complement whose subject is a (potentially empty) list of gap-synsems, as in the entry below:

(53) Subcategorization of 'think'

$$\begin{bmatrix} \mathit{word} \\ \textsc{phon}\ \mathit{think} \\ \textsc{ss loc cat val} \begin{bmatrix} \textsc{subj}\ \left\langle \text{NP} \right\rangle \\ \textsc{comps} \left\langle \begin{bmatrix} \textsc{loc cat} \begin{bmatrix} \textsc{head}\ \mathit{verb} \\ \textsc{val} \begin{bmatrix} \textsc{subj}\ \mathit{list(gap\text{-}ss)} \\ \textsc{comps}\ \langle\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

So, in a sentence like 'Who do you think pleased Robin?', 'think' will combine with a VP complement ('pleased Robin') which itself is slashed and whose subject is a gap-synsem. SLAC tells us that the SLASH value of any of the matrix verb's dependents should be amalgamated, so the SLASH value of the phrase 'pleased Robin' will become the SLASH value of the larger phrase 'think pleased Robin'. From there, things work as before, with the local value on SLASH eventually being bound off by a filler daughter with an identical local.

Adjunct extraction works pretty much the same way, and I won't discuss it here. The trick, of course, is to get the gap-synsem of the adjunct on the dependent's list of the head of the phrase it modifies. See Bouma et al. for a concrete suggestion on how this might be done.

3.3. **Extending the analysis to tough.** The theory of UDCs presented above extends quite straightforwardly to tough. In fact, Bouma et al. posit an entry like the one in (54) for this class of predicates.

(54) Bouma et al.'s lexical entry for 'easy'

$$\begin{bmatrix} \textsc{phon } \textit{easy} \\ \textsc{ss loc cat} \begin{bmatrix} \textsc{head } \textit{adj} \\ \textsc{arg-st} \begin{bmatrix} \textsc{subj} \left\langle \text{NP}_i \right\rangle \\ \textsc{comps} \left\langle \begin{bmatrix} \textsc{loc } \text{CP[to]} \\ \textsc{nonloc slash} \left\{ \boxed{1}\text{NP[acc]}_i \right\} \cup \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix} \\ \textsc{bind} \left\{ \boxed{1} \right\} \end{bmatrix} \end{bmatrix}$$

This entry is quite similar to the one proposed in Pollard & Sag 1994. That is, the subject of 'easy' is co-indexed with an accusative gap in the adjective's CP complement.[11] The gap is accusative, following Pollard & Sag 1994 in order to rule out embedded subject gaps, as in the ungrammatical (55).

(55) *Robin is easy to believe likes ice cream.

Unfortunately, requiring that the gap be accusative also seems to rule out the grammatical examples in (56).

(56) a. That Robin is a spy would be easy to imagine.
b. For him to like me is reasonable to expect.
c. Under the bed is hard for me to imagine being a reasonable place to keep your beer.
d. Proud of themselves is hard for me to picture ANY of them feeling, after what they've all done.

The co-indexation of the tough predicate's subject and the gap, as opposed to having fully identified local values as in a standard (i.e., "strong") UDC, is also a feature of the analysis in Pollard & Sag 1994, and exists in order to account for the case mismatch between filler and gap in examples like those in (57):

(57) a. He/*him is easy to please.
b. Who/*whom is easy to please.

[11] This complement should probably be a VP, as there is some evidence that the optional 'for'-phrase in which the subject of the downstairs verb is realized behaves like an argument of *easy*. Postal 1971, however, claims that the complement of these adjectives can be either S or VP, and that there is a subtle meaning difference between the two. I refer you to his book for more discussion.

Unfortunately, the co-indexing mechanism also apparently rules out the examples in (56), since the subjects of those sentences are not commonly thought of as bearing an index.

Otherwise, the entry is essentially fine. Note that SLAC ensures that the SLASH value corresponding to the tough-gap will **not** be amalgamated, while any other slash associated with the tough predicate's complement will. The reason is that the local value of the tough-gap is co-indexed with the BIND value of the tough predicate. The net effect of all this is that tough complements will not be islands for extraction, a correct prediction, as the following sentence illustrates:

(58) Which violins are these sonatas easy to play on?

What's left to do is adjust the theory in such a way that we are able to (a) allow the examples in (56), but (b) still rule out the ungrammatical examples in (55) and (57).

The following entry rules in the (56) examples, and retains the essential features of Bouma et al.'s proposal:

(59) Revised lexical entry for 'easy'

$$\begin{bmatrix} \textsc{phon} & \textit{easy} \\ \textsc{ss loc cat} & \begin{bmatrix} \textsc{head} & \textit{adj} \\ \textsc{arg-st} & \begin{bmatrix} \textsc{subj} & \left\langle \mathrm{XP}[\textsc{loc cont}\ \boxed{3}] \right\rangle \\ \textsc{comps} & \left\langle \begin{bmatrix} \textsc{loc} & \mathrm{VP[to]} \\ \textsc{nonloc slash} & \left\{ \boxed{1} \begin{bmatrix} \textsc{cont} & \boxed{3} \end{bmatrix} \right\} \cup \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix} \\ \textsc{bind} & \left\{ \boxed{1} \right\} \end{bmatrix} \end{bmatrix}$$

What has (substantively) changed here is that the content value on the slashed element corresponding to the tough-gap is now identified with the content value of easy's subject, as opposed to just an index. Preliminarily, no additional constraint is placed on the category or case of this gap.[12][13]

[12]Recall that the following, which are apparently counter-examples to the claim that non-NPs can be tough-subjects, are ruled out independently on semantic grounds.

(i): *Carefully is easy to imagine them fixing the car.

(ii): *Under the table is hard for me to look.

(iii): *To John is hard to give a book.

(iv): *There is hard to believe to be a unicorn in the garden.

[13]There is a subtle technical difficulty with this approach, as well. If the extracted element contains a quantifier (i.e., if it is a quantificational NP) then sharing the

We now need to say what it means for the tough gap and the tough subject to share the same CONTENT value. That is, under the assumptions made about the value of the feature CONTENT in Pollard & Sag 1994, in which an utterance's verb, verb phrase, and the sentence itself all shared the same content value, in accordance with the Semantics Principle, our entry for tough appears to incorrectly predict that the following, among others, would be grammatical.

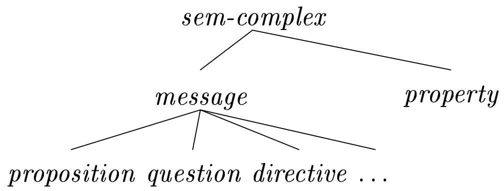
(60) a. *To prove that theorem is impossible for me to believe.
b. *That he be a spy is hard to make Robin.

In the case of (60a), the verb *believe* requires a sentential complement, and the subject is only a VP. But if the *content* value of VPs and sentences are identical, it is unclear how this is ruled out. A mismatch in the other direction occurs in (60b).

The solution to this problem involves adopting a finer-grained semantics of verbal elements, one which distinguishes semantically between verbs and verb phrases on the one hand, and sentences on the other. A proposal with exactly this feature has been made by Ginzburg & Sag 1998, motivated on entirely independent grounds.

Under Ginzburg and Sag's proposal, the value of the CONTENT feature of a verbal element will be a subtype of something called a *sem(antic)-complex*. The breakdown of *sem-complex* into subtypes is depicted below.

(61)



According to Ginzburg & Sag, the semantic content of a sentence will always be a subtype of *message*, whereas the semantic content of a verb or verb phrase will be a *property*. Returning to the previous examples, then, (60a) will be ruled out due to the fact that the semantic type of

CONTENT values in this way would result in two quantifiers appearing in different places in the semantics of the utterance. This is undesirable. The solution, it seems, involves changing the geometry of *content* values so that only the **nuclear** content (and not the quantificational content) of the tough subject and gap can be shared.

believe's complement will be required to be a *proposition*, whereas the subject of the utterance denotes a *property*, and similarly for (60b).[14]

Strikingly, this semantic ontology, coupled with the entry for tough above, also correctly predicts that the following will be grammatical. These examples can be thought of verbal analogues of the nominal "case mismatch" examples, cases which motivated Pollard & Sag 1994 to classify tough movement as a "weak" UDC.

(62) a. For Robin to be a spy would be hard to get over –.
b. *It would be hard to get over for Robin to be a spy.

(63) a. Him being a spy is hard to believe/prove.
b. *It is hard to believe/prove him being a spy.

(64) a. ?For him to be a spy is almost impossible to believe.
b. *It is almost impossible to believe for him to be a spy.

(65) a. For him to ignore the political consequences involved in an Atlantic Union of this kind is difficult to understand –.
b. *It is difficult to understand For him to ignore the political consequences involved in an Atlantic Union of this kind.[15]

Such examples seem quite grammatical, and are almost impossible to explain if full identification of the LOCAL value of the tough subject and its gap is required (as in a "strong" UDC). What I am advocating instead is a generalized version of the weak UDC account of Pollard & Sag 1994. The tough gap is an ordinary gap, and the construction behaves like a UDC with respect to parasitic gap licensing and island constraints. The "filler" in this construction, however, is actually just the BIND value associated with the tough predicate. The association of the gapped material with the tough subject is purely semantic, as evidenced by (a) case mismatches if the tough gap is nominal, and (b) verb form mismatches if the tough gap is verbal.

I turn now, and finally, to the issue of the ungrammaticality of (55). Any account of this somehow has to distinguish between subject and object gaps, which is not a completely straightforward task. In past UDC accounts of *tough*, some version of the following principle has been used, albeit stipulatively, to achieve the desired result:

[14]Note that trying to match a VP with a V would be ruled out by the constraint that the subject of tough be an XP, i.e., a COMPS saturated phrase.

[15]Again, I am grateful to Martin Jansche for this example. The grammatical example is from the Brown corpus.

The case value of an NP gap must be accusative.

The revised lexical entry for tough predicates, then, would simply incorporate that restriction into the selection of the gap.

(66) Lexical entry for ‘easy’ (predicative version)

$$\begin{bmatrix} \textsc{phon}\ \textit{easy} \\ \textsc{ss loc cat} \begin{bmatrix} \textsc{head}\ \textit{adj} \\ \textsc{arg-st} \begin{bmatrix} \textsc{subj} \left\langle \mathrm{XP}[\textsc{loc cont}\ \boxed{3}] \right\rangle \\ \textsc{comps} \left\langle \begin{bmatrix} \textsc{loc}\ \mathrm{VP}[\mathrm{to}] \\ \textsc{nonloc slash} \left\{ \boxed{1} \begin{bmatrix} \textsc{cont}\ \boxed{3} \end{bmatrix} \right\} \cup \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix} \\ \textsc{bind} \left\{ \boxed{1} \neg\ \mathrm{NP} \vee \textsc{ss loc cat head case}\ \textit{acc} \right\} \end{bmatrix} \end{bmatrix}$$

The intended reading of the entry above is admittedly a bit sloppy. What I really want is a condition on the value of SLASH which says that if the slash value is an NP, then it must be accusative. Such a condition is quite simple to write in RSRL (see Richter, 1998).

As I said, this account of the ungrammaticality of (55) is not very conceptually satisfying. It does illustrate, however, that it is possible within a UDC to account for this idiosyncratic property of the construction. Thus, I have provided an analysis of tough as an ordinary (albeit weak) UDC which addresses the only two real differences that emerged from our discussion of Grover and Cinque – (a) that the tough subject and the gap can differ in syntactic form (case or verb form), and (b) that tough gaps are prohibited in the subject position of finite clauses.

4. Further Issues

In §2, I used island sensitivity and parasitic-gap licensing as diagnostics for the UDC-hood of the tough construction, but I have not provided a theory of any of these effects. I take these to be largely independent issues, probably taking two separate papers to do justice. One might add crossover phenomena to these two cases if one were interested in a full theory of UDCs. In the case of crossover, there are many examples that are by now well-known, but for a which a satisfactory explanation has proved elusive.

There has been some activity on this front. For instance, Pollard 1996 and Sag 1998 are recent HPSG accounts of weak crossover which I believe to be consistent with the theory of tough presented here. Similarly, Bouma et al. 1998 suggest a direction for a theory of parasitic gaps.

On the issue of islands, the theory presented here already accounts for some island effects. For instance, the conjunct constraint, which prohibits moving one of the conjuncts in a coordinate structure, follows immediately from the traceless analysis presented here. Other island constraints, however, require more work. That is, the current theory does not immediately block extraction from an adverbial island, as in (67).

(67) *Which book did you file the report before anyone could read?

That is, nothing said so far prevents a slash introduced by a gap-synsem on the dependents list of 'read' from percolating by way of SLAC and SLIP to the SLASH of 'before', and thus to the dependents list of 'file', and so on.

Some past approaches to these examples have attempted to solve the problem by simply denying the existence of the effect. Counterexamples are trotted out, and the difficulty of examples like (67) are chalked up to extra-grammatical factors, like processing or pragmatics.

This approach seems misguided. Consider the NP islands discussed in §2, repeated here, and some well-known exceptions.

(68) a. *Who did Kim see a person who dislikes?
b. *Where did Kim like her brother from?

(69) a. Who did you take a photograph of?
b. Those guys are easy to paint a picture of?

The significance of the examples in (69) is controversial. Do they show that the grammar of English really does allow extraction from NPs? Or is there something special going on with picture NPs which allows them to violate an otherwise solid constraint? I lean toward the latter position. The alternative is to allow the grammar to license the utterances in (68). But they really seem "robustly" bad. That is, there is very little that can be done to improve them. In my opinion, it is this kind of data which constitutes the most important challenge to current HPSG-based theories of UDCs.

References

Ades, A. and M. Steedman (1982). On the order of words. *Linguistics and Philosophy 4*, 517–558.

Bouma, G., R. Malouf, and I. Sag (1998, January). Satisfying constraints on extraction and adjunction. Ms., Groningen University/Stanford University.

Calcagno, M., R. Levine, and T. Hukari (to appear). Parasitic gaps in English: some overlooked cases and their theoretical implications. In P. Culicover and P. Postal (Eds.), *Parasitic Gaps*. Cambridge, MA: MIT Press.

Chomsky, N. (1977). On Wh-movement. In P. Culicover, T. Wasow, and A. Akmajian (Eds.), *Formal Syntax*, pp. 71–132. New York: Academic Press.

Chomsky, N. (1986). *Barriers*. Cambrige, MA: MIT Press.

Cinque, G. (1990). *Types of $\overline{\mathrm{A}}$ Dependencies*. Cambridge, MA: MIT Press.

Fodor, J. D. (1983). Phrase structure parsing and the island constraints. *Linguistics and Philosophy 6*, 163–223.

Gazdar, G., E. Klein, G. Pullum, and I. Sag (1985). *Generalized Phrase Structure Grammar*. London: Blackwell.

Ginzburg, J. and I. Sag (1998). English interrogative constructions. Unpublished book, Hebrew University and Stanford University.

Grover, C. (1995). *Rethinking Some Empty Categories: Missing Objects and Parasitic Gaps in HPSG*. Ph. D. thesis, University of Essex.

Hukari, T. and R. Levine (1987). Rethinking connectivity in unbounded dependency constructions. In *Proceedings of the Sixth West Coast Conference on Formal Linguistics*.

Hukari, T. and R. Levine (1995). Adjunct extraction. *Journal of Linguistics 31*(2), 195–226.

Hukari, T. and R. Levine (1996). Phrase structure grammar: the next generation. *Journal of Linguistics 32*, 465–496.

Kaplan, R. and A. Zaenen (1989). Long-distance dependencies, constituent structure and functional uncertainty. In M. Baltin and A. Kroch (Eds.), *Alternative Conceptions of Phrase Structure*. Chicago: University of Chicago Press.

Kayne, R. (1983). Connectedness. *Linguistic Inquiry 14*, 223–249.

Kayne, R. and J.-Y. Pollock (1978). Stylistic inversion, succesive cyclicity, and move NP in French. *Linguistic Inquiry 12*, 93–133.

Kluender, R. (1998). On the distinction between strong and weak islands: a processing perspective. In P. Culicover and L. McNally (Eds.), *The Limits of Syntax*. New York: Academic Press.

Maling, J. and A. Zaenen (1978). The non-universality of a surface filter. *Linguistic Inquiry 9*, 475–497.

Manning, C. and I. Sag (1995). Dissociations between argument structure and grammatical relations. Ms., Stanford University.

McCloskey, J. (1979). *Transformational Syntax and Model-Theoretic Semantics*. Dordrecht: Reidel.

Morrill, G. (1994). *Type Logical Grammar: Categorial Logic of Signs*. Dordrecht: Kluwer.

Pickering, M. and G. Barry (1991). Sentence processing without empty categories. *Language and Cognitive Processes 6*(3), 229–259.

Pollard, C. (1996). Linear and hierarchical factors in weak crossover. Ms, Ohio State University.

Pollard, C. and I. Sag (1994). *Head-driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

Pollard, C. and E.-J. Yoo (1998). Quantifiers, *wh*-phrases and a theory of argument selection. *Journal of Linguistics*. (in press).

Postal, P. (1971). *Cross-Over Phenomena*. The Transatlantic Series in Linguistics. New York: Holt-Rinehart-Winston.

Przepiórkowski, A. (1996). Non-configurational case assignment. Paper presented at HPSG96, Marseille; published in Webelhuth et al. 1999, *Lexical and Constructional Aspects of Linguistic Explanation*, Stanford: CSLI.

Richter, F. (1998). RSRL for HPSG. In V. Kordoni (Ed.), *Tübingen Studies in HPSG*, Arbeitspapiere des SFB 340. Universität Tübingen.

Rosenbaum, P. (1967). *The Grammar of English Predicate Complement Constructions*. Cambridge, MA: MIT Press.

Sag, I. (1997). English relative clause constructions. *Journal of Linguistics 33*, 431–484.

Sag, I. (1998). Not a trace. Ms., Stanford University.

Sag, I. and J. D. Fodor (1994). Extraction without traces. In *Proceedings of the Thirteenth West Coast Conference on Formal Linguistics*.

Steedman, M. (1996). *Surface Structure and Interpretation*. Cambridge, MA: MIT Press.

Suñer, M. (1982). Big PRO and little *pro*. Ms., Cornell University.

Zwicky, A. (1986). The unaccented pronoun constraint in English. *Ohio State University Working Papers in Linguistics 42*.

Mike Calcagno

Seminar für Sprachwissenschaft
SFB 340, B8
Abt. Computerlinguistik
Kleine Wilhelmstr. 113
72074 Tübingen
Germany

E-mail address: calcagno@sfs.nphil.uni-tuebingen.de
Homepage: http://www.sfs.nphil.uni-tuebingen.de/~calcagno/