Satisfying Constraints on Extraction and Adjunction *

Gosse Bouma (gosse@let.rug.nl) and Robert Malouf (malouf@let.rug.nl)
University of Groningen

Ivan A. Sag (sag@csli.stanford.edu)
Stanford University

Abstract. In this paper, we present a unified feature-based theory of complement, adjunct, and subject extraction, in which there is no need either for valence reducing lexical rules or for phonologically null traces. Our analysis rests on the assumption that the mapping between argument structure and valence is defined by realization constraints which are satisfied by all lexical heads. Arguments can be realized as local dependents, in which case they are selected via the head's valence features. Alternatively, arguments may be realized in a long-distance dependency construction, in which case they are selected via the head's SLASH features. Furthermore, we argue that in English post-verbal adjuncts, as well as complements, are syntactic dependents selected by the verb, thus providing a uniform analysis of complement and adjunct extraction. Finally, we provide an alternative treatment of subject extraction which is subsumed by our general analysis and offer a new account of the *that*-trace effect.

1. Introduction

Feature-based analyses of filler-gap dependencies of the sort pioneered by Gazdar (1981) make a number of strikingly correct cross-linguistic predictions. Most notable among these, given that information about the 'extracted' element is locally encoded throughout the extraction path, is the prediction that some natural language phenomena might be sensitive to extraction information, e.g. a phenomenon that occurs only on an extraction path, or which is conditioned by the grammatical category of a particular extracted element some distance away. By now there is considerable cross-linguistic evidence confirming this prediction. Languages which exhibit extraction-sensitive phenomena include Irish (McCloskey, 1989), Chamorro (Chung, 1998), Palauan

* We would like to acknowledge a special debt to Bob Levine, Carl Pollard, Bob Kasper and Mike Calcagno, who have been skeptical of many of the ideas that led up to the proposals made here. This paper is much improved thanks to our interactions with them. In addition, we thank Georgia Green, Adam Przepiórkowski, Carl Pollard, and Peter Sells for their reactions to an earlier version of this paper, and Tania Avgustinova, Ash Asudeh, Myong-hi Chai, Ann Copestake, Dan Flickinger, Detmar Meurers, Jason Merchant, Jeff Smith, Gertjan van Noord, and three anonymous reviewers for helpful suggestions. Finally, this research was conducted in part as part of CSLI's English Resource Grammar Online (LINGO) project, in conjunction with the Verbmobil project. In that connection, we gratefully acknowledge the support of the Bundesministerium für Bildung, Wissenschaft, Forschung, und Technologie (Project VERBMOBIL) and the National Science Foundation (grant number IRI-9612682).

(Georgopolous, 1985), Icelandic (Maling and Zaenen, 1978), Kikuyu (Clements, 1984), Ewe (Collins, 1997), Thompson Salish (Kroeber, 1997), Moore (Haik, 1990), French (Kayne and Pollock, 1978), Spanish (Torrego, 1984), and Yiddish (Diesing, 1990), with no doubt others yet to be discovered.

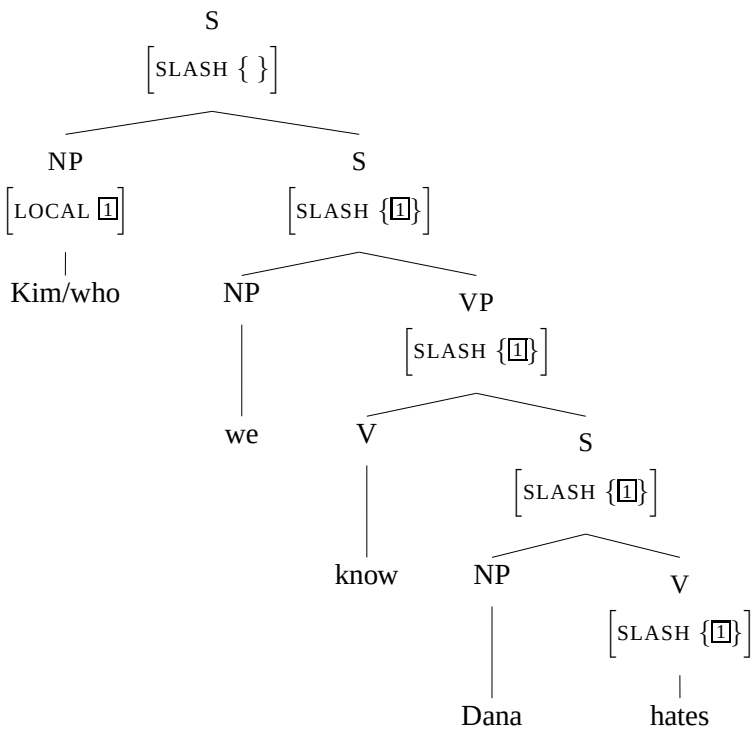
In a feature-based analysis, such phenomena are elegantly assimilated to local subcategorization. If phrases bear feature specifications indicating whether or not they contain a gap and, if so, what kind of element is missing from that phrase, then such information is locally accessible. Hence in just the same way that a word selects for some syntactic or semantic property of one of its arguments, it might select for an argument that is gap-free, an argument from which something has been extracted, or even for an argument containing a particular kind of gap.

Yet feature-based extraction analyses, admirable for their precision and comprehensiveness, have been less than fully satisfying. Specifically, work in Generalized Phrase Structure Grammar (Gazdar et al., 1985) and Head-driven Phrase Structure Grammar (e.g. Pollard and Sag, 1994; henceforth PS-94) has not yet been able to provide a unified account of extraction dependencies, relying instead on a number of unrelated mechanisms for complement, subject, and adverb extraction.

For example, PS-94 (chap. 9) posit a lexical rule that removes an element from a word's COMPS list, adding a compensating element to its SLASH value.[1] Such 'slashed' verbs combine locally with one fewer complement and pass their SLASH specification up to successively larger phrases. The slashed verb must thus be contained within a slashed phrase that combines with a compatible dislocated complement, as illustrated in (1).

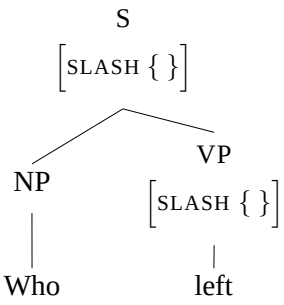
[1] Gazdar used the notation S/NP to indicate an S from which an NP element is extracted. SLASH has thus become the standard name for the set-valued feature encoding what elements are missing from a given constituent.

(1)

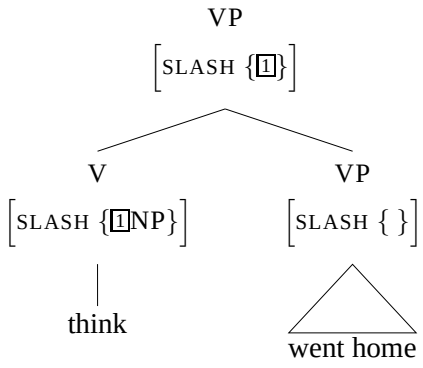


In contrast, subject extraction either is not treated as extraction at all, as in the slashless subject–VP analysis of (2a) (proposed originally by Gazdar (1981)), or else it involves a lexical rule which applies to the verb governing the clause from which the subject is extracted. The output of the lexical rule is a verbal sign specified as [SLASH {NP}] which selects for an unslashed VP, as in (2b).

(2) a.



b. (Who do you)



Finally, PS-94 treat examples like (3a) as combination of an adverbial modifier and an unslashed S, while the adverb in (3b) is the filler of an extraction construction whose lowest verb is slashed, courtesy of yet another lexical rule.

(3) a. Yesterday, we drank genever.

b. Yesterday, they think we drank genever.

The PS-94 account of English extraction thus involves three unrelated lexical rules and entails a somewhat unintuitive distinction between slashed and unslashed instances of subject or adverb extraction.

As noted by Hukari and Levine (1996a) in their review of PS-94, the resulting system is inelegant and seems to give the wrong predictions for certain instances of parasitic gaps. To this objection, Hukari and Levine (1996b) add the further observation that cross-linguistic evidence indicates more unity within extraction dependencies than is embodied in the PS-94 analysis. In languages where extraction is locally registered along the extraction path, extraction of embedded subjects is indicated just as extraction of other arguments is. For example, in Chamorro, where verbs bear morphology (indicated as WH in the following glosses) indicating agreement with arguments that either contain or are themselves extracted elements, we find that embedded subjects and main clause subjects trigger the very same morphological registration:[2]

(4) a. Hayi f-um-a'gasi i kareta
who WH.SU-wash the car

'Who washed the car?'

b. Hayi si Juan ha-sangan-i hao [f-um-a'gasi i kareta]
who UNM Juan E3S-say-DAT you WH.SU-wash the car

'Who did Juan tell you washed the car?'

[2] Our assumptions about Chamorro extraction are based on those of Chung (1982, 1994), which are challenged by Dukes (1992). For a reply to Dukes, see Chung (1998).

c. Hafa um-istotba hao [ni malagao'-na i
what WH.SU-disturb you COMP WH.OBL-want-3SG the
lahi-mu]
son-your

'What does it disturb you that your son wants?'

In all cases of *wh*-agreement we are aware of, extraction of main clause and embedded subjects is uniformly registered. The verb morphology demonstrated in (4) encodes the information 'my subject either is or contains a gap'. Facts like these suggest that the slashless, in situ analysis of matrix and embedded subjects, suggested originally by Gazdar and adopted by PS-94, is inconsistent with the elegant generalization that all *wh*-agreement phenomena can be treated in terms of selection for slashed elements by a lexical head.

Moreover, as argued at length by Hukari and Levine (1995), adverbial extraction is registered in many of these languages in just the same way as complement extraction. Yet if adverbials are adjoined modifiers that select syntactically for the phrases they modify (as is standardly assumed), then adverb extraction involves no verb whose argument is slashed. Unless some further ad hoc machinery is introduced, the simple generalization embodied in SLASH-based extraction analyses—namely that extraction is lexically registered as selection for a slashed argument—cannot be extended to the registration of adverb extraction. This is particularly problematic in the case of languages such as Chamorro, where *wh*-agreement is unquestionably lexical in nature.

In this paper, we offer a remedy for this apparent failing of SLASH-based extraction analyses. We present a novel, traceless account of complement extraction, in which there is no need for a Complement Extraction Lexical Rule. Rather, our analysis rests on the assumption (defended in Section 2) that the mapping between argument structure and valence is defined by general constraints which apply to both unslashed and slashed lexical signs. Section 3 introduces a constraint-based account of complement extraction which differs from previous feature-based accounts in that it eliminates the configurationally defined Foot/Non-local Feature Principle of Gazdar et al. (1985) and PS-94 in favor of a lexically based constraint. While the account we present here preserves features of Sag's (1997) traceless analysis of extraction, it differs in making no use of lexical rules to establish a relation between unslashed lexical entries and their slashed counterparts. In addition, we show how the present proposal handles certain constructions (such as parasitic gaps and *tough* movement) that have figured prominently in earlier discussions. We also review some of the evidence against the existence of *wh*-traces.

Next, in section 4 we turn to adjuncts. There is now a significant literature providing cross-linguistic evidence that, contrary to traditional wisdom,

many adjuncts must be analyzed not as selecting syntactically for the phrase they modify semantically, but rather as being selected by a lexical head in much the same way that complements are. We follow this line of research, and argue that in English, post-verbal adjuncts as well as complements are syntactic dependents selected by the verb. As shown in section 5, this not only offers a straightforward account of the fact that adjuncts may be extracted, it also predicts Hukari and Levine's observation that complement and adjunct extraction is registered uniformly in a wide range of languages.

Finally, we provide an alternative account of subject extraction which is subsumed by our general theory of extraction. Our account of the *that*-trace effect does not use anything like the Subject Extraction Lexical Rule of PS-94. It also solves the problem (observed by Hukari and Levine, 1996a) that extracted subjects may license parasitic gaps, a fact that is inconsistent with proposals involving a Subject Extraction Lexical Rule (see above), which treat sentences with extracted subjects as plain, gap-free VPs.

2. Background Assumptions

In this section, we introduce the theoretical concepts and assumptions on which we will build our analysis of adjunction and extraction.

Foremost is the distinction between syntactic valence and selected arguments, which will play a key role in the analysis of adjunction and extraction. Selected arguments, we argue, are the elements which a lexical item truly subcategorizes for. Following earlier work in HPSG, we represent this level as a list-valued feature which we will call, following Manning and Sag (1998), ARGUMENT-STRUCTURE (ARG-ST). This is distinct from the valence features (SUBJ, COMPS, and SPR), which represent the local dependents of a lexical head. Syntactic cancellation, as defined by the Valence Principle, is relevant only for the three valence features. The relationship between selected arguments and local dependents is expressed by means of an intermediate level of representation, a kind of extended argument structure that we treat via a distinct list-valued feature DEPENDENTS (DEPS). Argument structure plays a crucial role in the definition of binding theory; the dependents list is fundamental to the traceless analysis of extraction presented below. [3]

Second, we will assume that the subcategorization properties of a lexical item underspecify its valence properties. That is, in addition to the canonical case, where each element on the ARG-ST list corresponds to an element on

[3] This is not to say, however, that a treatment of adjuncts as complements and adjunct extraction has to make use of DEPS. In Przepiórkowski (1999a) an analysis of adjuncts as complements and adjunct extraction is presented which is to a large extent compatible with ours but which follows Manning et al. (1999) in that it assumes a lexical rule for adding adjuncts to ARG-ST.

one of the valence lists and vice versa, we also allow for cases where the relationship between ARG-ST and valence is not one to one. We will deal with two such cases in detail below: (1) extraction is analyzed as the case where a member of the dependency structure (the extended argument structure) of a lexical head is not realized on any valence list and hence not realized as a local dependent; and (2) selection of adjuncts results when an element absent from a head's argument structure appears both on its DEPS list and its COMPS list. The various possible associations between ARG-ST and the valence features, mediated by the DEPS list, are expressed by constraints on types in the lexical hierarchy. The introduction of such constraints provides a declarative alternative to lexical rules which is compatible with lexical architectures based on inheritance.

2.1. Selected Arguments, Argument Structure, and Valence

In previous versions of HPSG (Pollard and Sag, 1987; the first eight chapters of Pollard and Sag, 1994), the feature SUBCAT played a dual role: it characterized a lexical head's *argument structure* as well as its *valence*, i.e. its local combinatoric potential. In more recent work (Borsley, 1989a, 1989b; chapter 9 of Pollard and Sag, 1994; and especially Manning and Sag, 1998), it has been argued that these two functions should be performed by distinct features. *Valence* has been reanalyzed in terms of the features SUBJ, SPR, and COMPS, while the feature ARG-ST has been introduced to define the *argument structure* of lexical items and to provide the appropriate level of representation for binding theory. In this paper, we argue that both of these levels, together with a third level of dependency structure (represented by the DEPS feature mentioned above) play a role in the grammar of adjunction and extraction. Arguments for the distinction between argument structure and valence are given below. In the next section we will motivate the introduction of DEPS, and explain how the relationship between the three different levels of representation can be defined succinctly by means of relational constraints.

Valence features are defined for words as well as phrases, and define the local combinatoric potential of these elements. Argument structure, on the other hand, is only defined for words and is the level to which the binding constraints apply (taking into account information about the arguments that are 'unified in' through syntactic combination). Valence features are 'canceled' in syntax; as phrasal nodes do not have a value for ARG-ST, it does not make sense to define any form of cancellation for this feature.

Valence and argument structure are clearly related, and for words, which have both argument structure and valence features, it is tempting to simply conflate the two notions. While this may be possible in many cases, there are nevertheless a number of reasons to distinguish the two.

First, the distinction between the 'syntactic' subject and the 'underlying' or 'logical' subject, which is relevant for the analysis of passive and ergativity in a number of languages, requires that valence and argument structure be represented separately (Manning, 1996; Manning and Sag, 1998). For instance, in a language such as Toba Batak, a single argument structure can be realized using active or objective voice:

(5) a. Mang-ida si Ria si Torus
AV-see PM Ria PM Torus

'Torus sees/saw Ria.'

b. Di-ida si Torus si Ria
OV-see PM Torus PM Ria

'Torus sees/saw Ria.'

As argued extensively by Kroeger (1993), the NP immediately following the verb is syntactically an object, forming a VP with the verb. This can be accounted for if there is a single lexically specified argument structure which corresponds to two possible realizations of valence, one in which the first element on ARG-ST is realized as the syntactic subject, and one in which the second is. This analysis, together with the assumption that binding theory applies to argument structure, accounts for the data below.

(6) a. [Mang-ida diri-na$_i$] si Torus$_i$
AV-see self-his PM Torus

'Torus$_i$ sees/saw himself$_i$.'

b. [Di-ida si Torus$_i$] diri-na$_i$
OV-see PM Torus self-his

'Torus$_i$ sees/saw himself$_i$.'

In (6a), the first argument *Torus* is realized as the subject and the second argument *dirina* 'himself' is realized as the direct object. In (6b), this is reversed: *dirina* is realized as the subject and *Torus* is realized as the direct object. Regardless of how the arguments are realized, however, the first element of the argument structure is always the antecedent of the reflexive second argument.

Another reason to keep argument structure separate from valence is that such a distinction also seems important in providing an account of so-called '*pro*-drop' phenomena in languages such as Spanish, where arguments which are syntactically unexpressed nonetheless play a grammatically significant role:

(7) a. llegué
pro arrive-1.SG
'I arrived.'

b. llegó
pro arrive-3.SG
'He/she arrived.'

Evidence of a different kind for distinguishing argument structure and valence comes from the literature on complement inheritance. Complement inheritance (or argument composition) has been used to account for the word order of German and Dutch verb complexes (Hinrichs and Nakazawa, 1994; van Noord and Bouma, 1996), French clitic climbing (Miller and Sag, 1997), and Japanese causatives (Manning et al., 1999), among others. As observed by van Noord and Bouma (1996), complement inheritance is incompatible with the notion of 'local domain' which is crucial for binding theory. If binding applies to argument structure, however, and complement inheritance is defined for the valence feature COMPS only, the problem is avoided.

Finally, an essential property of our traceless extraction analysis is that it must admit lexical signs in which only a subset of the elements on ARG-ST are realized via the valence features. However, reconstruction effects show that binding theory applies to all selected elements of a lexical head, irrespective of extraction. Thus, there must be a level of represetation (viz. argument structure) where these are all visible.

We conclude that there are both theory independent considerations (ergativity and *pro*-drop) and theory internal considerations (the elimination of traces and the introduction of complement inheritance) which suggest that the binding theory as formulated in PS-94 must apply to a level of representation which is distinct from syntactic valence. This motivates the introduction of ARG-ST as a level of representation which cannot be reduced to valence. It also raises the issue of how the close relationship between these two levels is best accounted for.

2.2. A CONSTRAINT-BASED ACCOUNT OF VALENCE

In this section, we argue that the introduction of an additional level of representation need not give rise to redundancy. Furthermore, distinguishing between argument structure and valence has the important additional benefit that variations in valence of a given lexical head, which previously were accounted for by means of lexical rules, can now be accounted for by means of constraints on the relationship between (extended) argument structure and valence.

For non-ergative, non *pro*-drop languages such as English, one might suggest that argument structure can be derived from valence by means of the constraint in (8).[4]

(8) $$\textit{word} \Rightarrow \begin{bmatrix} \text{SUBJ} & \boxed{1} \\ \text{COMPS} & \boxed{2} \\ \text{ARG-ST} & \boxed{1} \oplus \boxed{2} \end{bmatrix}$$

That is, the valence features of a 'canonical' lexical entry add up[5] to argument structure. Redundancy between the two levels of representation is avoided, since a lexical entry only needs to specify its valence properties or its argument structure. In the latter case, the values of the features SUBJ and COMPS follow, as long as one assumes that independent constraints will require the SUBJ value of a verb to have exactly one element, the SUBJ value of a case-marking preposition will be empty, etc.

Lexical entries which do not follow this pattern could be derived by means of lexical rules. Two lexical rules producing such 'non-canonical' lexical entries are the Complement Extraction Lexical Rule of Sag (1997) and the Adjunct Lexical Rule of van Noord and Bouma (1994) and Manning et al. (1999). Schematic versions of these rules are given in (9).

(9) a. $$\begin{bmatrix} \textit{word} & \\ \text{SUBJ} & \boxed{1} \\ \text{COMPS} & \boxed{2} \end{bmatrix} \mapsto \begin{bmatrix} \text{SUBJ} & \boxed{1} \\ \text{COMPS} & \boxed{2} \ominus \langle \textit{gap-ss} \rangle \end{bmatrix}$$

b. $$\begin{bmatrix} \textit{verb} & \\ \text{SUBJ} & \boxed{1} \\ \text{COMPS} & \boxed{2} \end{bmatrix} \mapsto \begin{bmatrix} \text{SUBJ} & \boxed{1} \\ \text{COMPS} & \boxed{2} \oplus \langle \text{'adverbial'} \rangle \end{bmatrix}$$

The Complement Extraction Lexical Rule in (9a) removes an element from COMPS and instantiates it as a 'gap' (i.e. an element of type *gap-synsem* (*gap-ss*)). The Adjunct Lexical Rule in (9b) adds an adverbial *synsem* to COMPS.[6]

[4] We will ignore the feature SPR from now on, as it plays no role in the remainder of the paper.

[5] We use $A \oplus B$ to denote the concatenation (or 'append') of the lists A and B. and $A \ominus B$ to denote the list A minus the elements in B. For present purposes, $\ominus$ can be thought of as interdefinable with the domain union operator ($\bigcirc$) of Reape (1994) and Kathol (1995):

$$(A \ominus B = C) \Leftrightarrow (C \bigcirc B = A)$$

[6] Basically, a *synsem* is a potential *adverbial* if its MOD feature is unifiable with the *synsem* value of the head it modifies (i.e. the item on whose COMPS list it appears). Semantically, the extra element acts as a modifier of the head. See van Noord and Bouma (1994), Manning et al. (1999), and section 4 for details.

In both cases, the resulting lexical entries exhibit a mismatch between argument structure and valence.

In the following sections we will discuss the motivation for a traceless account of extraction and for an account of adjunction in which adjuncts are selected by the same mechanism that is responsible for selection of complements. At this point, however, we want to draw attention to the fact that both proposals as implemented in (9) use lexical rules to derive lexical entries which do not satisfy the constraint in (8). Note, however, that while both lexical rules nonmonotonically alter the valence properties of a lexical entry, they only monotonically alter argument structure (the CELR instantiates an argument as *gap-ss* and the Adjunct LR leaves argument structure unaffected). This suggests that the effect of such lexical rules can also be achieved by reformulating the constraint in (8) so as to allow both canonical and non-canonical relationships between argument structure and valence. This is the approach we explore below.

First of all, the introduction of adjuncts on COMPS blurs the distinction between adjuncts and truly selected dependents. The distinction is relevant, if only because a lexical item needs to specify the elements for which it selects. To preserve this distinction, we will therefore assume first the level of ARG-ST, which contains all and only the selected arguments of a lexical head. In addition, we introduce dependency structure as an extended argument structure. The feature DEPS specifies the list of dependents of a lexical head. In the case of verbs, this consists of the selected arguments plus an underspecified list of adverbial *synsem*s. The relationship between ARG-ST and DEPS is therefore defined by means of the following constraint:

(10) **Argument Structure Extension**:

$$\textit{verb} \Rightarrow \begin{bmatrix} \textsc{arg-st} & \boxed{1} \\ \textsc{deps} & \boxed{1} \oplus \textit{list}(\textit{`adverbial'}) \end{bmatrix}$$

Here we are concerned mainly with adjuncts of verbal heads, so the Argument Structure Extension is assumed to be restricted to verbs. Incorporation of adjuncts in the argument structure of nouns is used in the analysis of West Greenlandic deverbal nouns presented in Malouf (1999).

Second, the traceless account of extraction employs a lexical rule for removing elements from COMPS. Our constraint-based alternative allows one or more *gap-ss* elements to be present on DEPS, but not on COMPS:

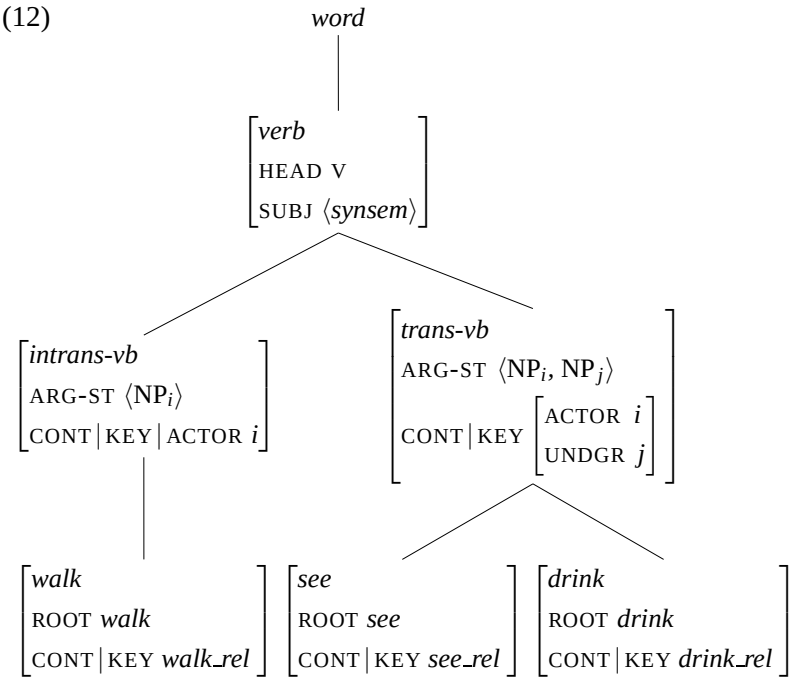
(11) **Argument Realization**:

$$\textit{word} \Rightarrow \begin{bmatrix} \textsc{subj} & \boxed{1} \\ \textsc{comps} & \boxed{2} \ominus \textit{list}(\textit{gap-ss}) \\ \textsc{deps} & \boxed{1} \oplus \boxed{2} \end{bmatrix}$$

Argument Realization defines the relationship between dependents and valence, but differs from the constraint given earlier in (8) in that it allows (non-subject) elements on DEPS that are of type *gap-ss* to be absent from the COMPS list. In fact, because each member of the COMPS list must be expressed as a *sign*, it follows from the fact that all *sign*s have a canonical SYNSEM value (see the constraint given below in (28)) that all non-subject *gap-ss* elements on the DEPS list are required to be absent from the COMPS list.

The constraints in (10) and (11) replace the lexical rules in (9). Since they are constraints on lexical types, they impose constraints on (subparts of) the lexicon. For instance, a simplified definition of the lexical type *verb* takes the form of the partial hierarchy in (12).

(12)



Verbs are a subtype of *word* whose HEAD value is of type *v* and which select for a subject. Subtypes of *verb* define various subcategorization types by providing a value for ARG-ST, as well as a definition of how these arguments contribute to the semantic content of the verb. The most specific types define individual lexemes, for which only ROOT and CONT values need to be specified.[7]

[7] It should be noted that the view of the lexicon presented above is in certain ways an oversimplification. The relationship between argument structure and semantics in particular is

All verbal lexical signs must also satisfy Argument Structure Extension and Argument Realization. If we add this information, the 'enriched' lexical entry for a transitive verb such as *drink*, whose ARG-ST consists of two NP *synsem*s, therefore is as follows:

(13)
$$\begin{bmatrix} \textit{verb} & \\ \textsc{root} & \textit{drink} \\ \textsc{subj} & \langle \boxed{1} \rangle \\ \textsc{comps} & \boxed{3} \ominus \textit{list}(\textit{gap-ss}) \\ \textsc{deps} & \langle \boxed{1} \rangle \oplus \boxed{3}(\langle \boxed{2} \rangle \oplus \textit{list}(\text{`}\textit{adverbial}\text{'})) \\ \textsc{arg-st} & \langle \boxed{1}\text{NP}_i, \boxed{2}\text{NP}_j \rangle \\ \textsc{cont}|\textsc{key} & \begin{bmatrix} \textit{drink\_rel} \\ \textsc{actor}\ i \\ \textsc{undgr}\ j \end{bmatrix} \end{bmatrix}$$

As both the list of *gap-ss* elements and the list of adverbials is underspecified in the lexicon, this description can be satisfied by many distinct feature structures. Three possibilities are considered below. If both lists are empty, the description simplifies to (14). Feature structures satisfying this description must have a COMPS value containing only a direct object NP.

(14)
$$\begin{bmatrix} \textit{verb} & \\ \textsc{root} & \textit{drink} \\ \textsc{subj} & \langle \boxed{1}\text{NP}_i \rangle \\ \textsc{comps} & \langle \boxed{2}\text{NP}_j \rangle \\ \textsc{deps} & \langle \boxed{1}, \boxed{2} \rangle \\ \textsc{arg-st} & \langle \boxed{1}, \boxed{2} \rangle \\ \textsc{cont}|\textsc{key} & \begin{bmatrix} \textit{drink\_rel} \\ \textsc{actor}\ i \\ \textsc{undgr}\ j \end{bmatrix} \end{bmatrix}$$

If the list of adverbials contains a single element and the list of gaps is empty, the lexical description simplifies to (15).

stipulated here in the types for intransitive and transitive verbs. But, as many have observed, there are important regularities to be captured in this domain. Such observations are not at all incompatible with the view presented above; indeed, they can be incorporated by introducing a more fine grained approach to subtyping of verbs. For recent HPSG analyses along these lines see Wechsler (1995), Davis (1996), and Koenig (1999).

(15)
$$\begin{bmatrix} \textit{verb} & \\ \textsc{root} & \textit{drink} \\ \textsc{subj} & \langle \boxed{1}\text{NP}_i \rangle \\ \textsc{comps} & \langle \boxed{2}\text{NP}_j, \boxed{3}\text{`}\textit{adverbial}\text{'} \rangle \\ \textsc{deps} & \langle \boxed{1}, \boxed{2}, \boxed{3} \rangle \\ \textsc{arg-st} & \langle \boxed{1}, \boxed{2} \rangle \\ \textsc{cont}|\textsc{key} & \begin{bmatrix} \textit{drink\_rel} \\ \textsc{actor} \ i \\ \textsc{undgr} \ j \end{bmatrix} \end{bmatrix}$$

Feature structures satisfying this description must have a COMPS value with two elements. Finally, if the list of gaps contains the element corresponding to the direct object, and the list of adverbials is empty, we obtain the simplified description in (16).

(16)
$$\begin{bmatrix} \textit{verb} & \\ \textsc{root} & \textit{drink} \\ \textsc{subj} & \langle \boxed{1}\text{NP}_i \rangle \\ \textsc{comps} & \langle\ \rangle \\ \textsc{deps} & \langle \boxed{1}, \boxed{2}\text{NP}[\textit{gap-ss}]_j \rangle \\ \textsc{arg-st} & \langle \boxed{1}, \boxed{2} \rangle \\ \textsc{cont}|\textsc{key} & \begin{bmatrix} \textit{drink\_rel} \\ \textsc{actor} \ i \\ \textsc{undgr} \ j \end{bmatrix} \end{bmatrix}$$

In this case, only feature structures whose COMPS list is empty satisfy the description.

Note that in the kind of architecture presented here, it makes sense to distinguish between lexemes and full blown feature structure descriptions of type *word*. Lexemes are the descriptions at the bottom of the lexical hierarchy, specifying only a value for ROOT and CONT. These descriptions are enriched by inheritance, to give rise to more complex feature descriptions of the form shown in (13). Each lexical entry corresponds to exactly one enriched lexical description. Thus, lexical inheritance allows us to define complex lexical descriptions without redundancy.

Feature structure descriptions, however, only define a set of constraints on the feature structures satisfying these descriptions, and thus in general a single description can be satisfied by many different feature structures. This is true in particular for the lexical descriptions in our theory. Argument Structure Extension allows DEPS to contain an arbitrary number of elements in addition to the selected arguments found on ARG-ST. Therefore, a single lexical entry,

specifying a fixed number of arguments on ARG-ST, can be satisfied by an infinite number of feature structures, differing only in the number of elements they contain on DEPS. This method is no different in principle from that standardly assumed in lexical description. For example, a lexical description requiring that a verb's object be a noun phrase of (say) genitive case allows there to be infinitely many phrases that can serve as the given verb's object. These phrases may differ wildly in content, person, number, gender, and so forth, just as long as they satisfy the lexically specified constraint requiring them to be genitive NPs.

The infinity which is a consequence of Argument Structure Extension is also similar to the infinity which arises as a consequence of recursive lexical rules (i.e. rules which may apply to their own output). For example, the Adjunct Lexical Rule allows a single lexical item to give rise to an infinite number of derived items. As argued in van Noord and Bouma (1994), the computational problem posed by this kind of recursion can be solved by reformulating lexical rules as recursive constraints on lexical entries, whose evaluation can be delayed to a point where only a finite number of solutions remain (typically, after some syntactic processing has taken place). Of course, if lexical rules are to be interpreted as constraints, a more perspicuous analysis results by specifying the relevant constraints (e.g. Argument Structure Extension and Argument Realization) directly and eliminating the lexical rules from the grammar altogether.

2.3. REMAINING ISSUES

In the examples above we have left out many details of lexical entries. In particular we have suppressed all paths dominating the valence features, ARG-ST, and DEPS. A more complete picture of the architecture of signs of type *word* is given in (17). This architecture differs from that of previous presentations of HPSG in that it distinguishes a new feature DEPS both from the valence features and from ARG-ST. The features DEPS and ARG-ST are relevant for words only, and thus do not appear on signs of type *phrase*. Note also that the feature ROOT takes a *morpheme* as its value, and I-FORM's value is an *inflected form*. The inflected form of a word is normally identical to its PHON value (but see Koenig (1994), Miller and Sag (1997), and Abeillé et al. (1998) for accounts of French pronominal cliticization in which the relationship between PHON and I-FORM is more complex).

(17)
$$\begin{bmatrix} \textit{word} \\ \text{PHON} \quad \textit{phon-form} \\ \text{MORPH} \quad \begin{bmatrix} \text{ROOT} \quad \textit{morph} \\ \text{I-FORM} \ \textit{infl-form} \end{bmatrix} \\ \text{SYNSEM} \quad \begin{bmatrix} \textit{synsem} \\ \text{LOC(AL)} \quad \begin{bmatrix} \textit{loc} \\ \text{CAT} \quad \begin{bmatrix} \text{HEAD} \quad \begin{bmatrix} \text{MOD} \ \textit{synsem} \end{bmatrix} \\ \text{VAL} \quad \begin{bmatrix} \text{SUBJ} \quad \textit{list}(\textit{synsem}) \\ \text{SPR} \quad \textit{list}(\textit{synsem}) \\ \text{COMPS} \ \textit{list}(\textit{synsem}) \end{bmatrix} \\ \text{ARG-ST} \ \textit{list}(\textit{synsem}) \\ \text{DEPS} \quad \textit{list}(\textit{synsem}) \end{bmatrix} \\ \text{CONT} \ \textit{cont} \end{bmatrix} \\ \text{NONLOC} \ \begin{bmatrix} \text{SLASH} \ \textit{set}(\textit{loc}) \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

The phrase structure schemata used in what follows are the Head-Subject and Head-Complement Schemata familiar from PS-94. The Head Feature Principle, which identifies the values of the feature HEAD on the mother and head daughter, and the Valence Principle, which defines the value of the SUBJ, COMPS, and SPR on the mother to be the result of subtracting the non-head daughters' *synsem*s from the corresponding valence list on the head-daughter, apply to these schemata as usual.

However, as explained in detail in the next section, our account of extraction differs from that of PS-94. In particular, there is no *phrasal* constraint analogous to the Nonlocal Feature Principle, which defined a mother's SLASH value as the union of the daughters' SLASH values.

We also depart from earlier work in our assumptions concerning semantics. The semantics principle of PS-94 picks out one of the daughters of a phrase as the semantic head (the adjunct daughter if present, the head daughter otherwise), which supplies the semantics of the phrase as a whole. The semantic head in turn is responsible for incorporating the semantics of the other daughters by unification. Below, we will adopt the semantic architecture of *minimal recursion semantics* (Copestake et al., 1999), in which the semantics of a phrase is the union of the sets of semantic constraints and relations provided by the daughters. Thus, for headed phrases we assume the following semantic composition principle:

(18) **Semantic Composition:**

headed-ph ⇒

$$\left[\begin{array}{ll} \textsc{cont} & \left[\begin{array}{l}\textsc{key}\ \boxed{1}\\ \textsc{rels}\ \boxed{2}\oplus\boxed{3}\oplus\cdots\oplus\boxed{i}\end{array}\right] \\ \textsc{head-dtr} & \left[\textsc{cont}\left[\begin{array}{l}\textsc{key}\ \boxed{1}\\ \textsc{rels}\ \boxed{2}\end{array}\right]\right] \\ \textsc{non-head-dtrs} & \left\langle\left[\textsc{cont}\,|\,\textsc{rels}\ \boxed{3}\right],\ldots,\left[\textsc{cont}\,|\,\textsc{rels}\ \boxed{i}\right]\right\rangle \end{array}\right]$$

Here RELS (relations) is a list of elementary predications. The value of the feature KEY is the semantic relation introduced by the lexical head of the phrase and is passed up from the head of a phrase to the mother. For words, we may assume that in general their semantics coincides with the KEY semantic relation they introduce:

(19) *word* ⇒ $\left[\textsc{cont}\left[\begin{array}{l}\textsc{key}\ \boxed{1}\\ \textsc{rels}\ \langle\boxed{1}\rangle\end{array}\right]\right]$

Although semantic issues play only a very minor role in what follows, this conception of semantic composition is crucial for our constraint-based treatment of adjuncts as dependents.

3. Complement Extraction

In this section, we develop a novel approach to extraction. It differs from previous feature-based accounts in that it is traceless, head-driven, and involves no lexical rule of the sort proposed in PS-94, Sag and Fodor (1994), or Sag (1997). By 'head-driven' we mean that a phrase inherits SLASH information only from its head daughter, rather than inheriting from all of its daughters as it does in analyses based on the Foot Feature Principle of Gazdar et al. (1985) or the Nonlocal Feature Principle of PS-94.

The approach to extraction developed here has considerable overlap with the proposal made in Sag (1997), but it makes one minor and one major modification. The minor change is that in order to allow for the account of adjunction presented in section 4, SLASH amalgamation is defined for DEPS rather than ARG-ST. The major modification is that the Complement Extraction Lexical Rule is eliminated, as are the lexical rules which have been proposed for subject and adjunct extraction. Instead, Argument Realization will account for the lexical variation required for a traceless account of

extraction. This has at least one highly significant consequence: it provides an entirely uniform account of complement, subject, and adjunct extraction. In addition, it eliminates a lexical rule whose status has always been dubious and whose interaction with other analyses has been highly controversial.

The following section introduces the technical details of our analysis. We then discuss the consequences of the current proposal for the analysis of some of the languages mentioned in the introduction, for *easy* constructions, and for cases where there is a mismatch between syntactic configuration and the locations of gap introduction and binding. Our proposal follows a tradition within nontransformational grammar (including, among others, Gazdar et al. (1984), Steedman (1996), Morrill (1994), and Kaplan and Zaenen (1989)) demonstrating the feasibility of traceless accounts of extraction. We conclude this section with a critical assessment of the most compelling arguments that have been presented for and against the existence of *wh*-traces.

3.1. CONSTRAINT-BASED HEAD-DRIVEN EXTRACTION

HPSG analyses of extraction, building on earlier work in GSPG, involve values of the feature SLASH that are projected upward in a syntactic structure as illustrated in (1) above. A non-empty SLASH value on a phrasal sign signals that an element is 'missing' from the corresponding phrase. The distribution of SLASH normally follows from its interpretation as a placeholder for missing elements: if a node dominates a daughter containing a non-empty SLASH value, the SLASH value of the node itself must also be nonempty. Two kinds of exceptions to this rule are possible: in head-filler constructions (i.e. the top of a *wh*-question or of a relative clause) a filler combines with a head whose SLASH value unifies with (the LOCAL value of) the filler. The missing element is 'found' at this point, and thus the SLASH of the resulting phrase is empty. A second type of exception may occur if a lexical head selects an incomplete complement, as do adjectives like *easy*, *tough*, and the like.

The distribution of SLASH is normally accounted for by the Nonlocal (or Foot) Feature Principle, which constrains the SLASH value on phrasal nodes to be the union of the SLASH values on its daughters. This approach is configurational in that it defines SLASH in terms of mothers and daughters in a phrase structure tree. Also, it implements a strict bottom-up flow of information. Nonempty SLASH values are introduced by traces and these values are consistently passed up to nodes higher in the tree. This scheme is satisfactory in the majority of cases, but, as we shall see below, there are at least some data which are hard to reconcile with it. Specifically, adjectives like *easy* 'bind' an element in the SLASH value of one of their dependents—their infinitival complement.

In our lexicalist alternative to configurational theories of SLASH, the Nonlocal Feature Principle is replaced by a simpler constraint on phrases and a

lexical constraint on heads.[8] More precisely, we propose that the SLASH value of a lexical item is defined in terms of the SLASH values of its dependents. Lexical binding of SLASH is accounted for by the feature BIND, which has the empty set as value for all words except SLASH binders like *easy* and *tough*:

(20) **SLASH Amalgamation**:

$$\textit{word} \Rightarrow \begin{bmatrix} \text{LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{DEPS} & \left\langle \left[\text{SLASH } \boxed{1}\right], \ldots, \left[\text{SLASH } \boxed{n}\right] \right\rangle \\ \text{BIND} & \boxed{0} \end{bmatrix} \end{bmatrix} \\ \text{SLASH} & \left(\boxed{1} \cup \ldots \cup \boxed{n}\right) - \boxed{0} \end{bmatrix}$$

SLASH Amalgamation ensures that if a dependent is slashed then the head which selects it will also be slashed. So, by this constraint, the verb *know* in (21) is slashed if either its subject or its sentential complement is slashed.

(21)
$$\begin{bmatrix} \text{I-FORM} & \textit{know} \\ \text{SUBJ} & \langle \boxed{3} \rangle \\ \text{COMPS} & \langle \boxed{4} \rangle \\ \text{DEPS} & \left\langle \boxed{3}\begin{bmatrix} \text{LOC} & \text{NP} \\ \text{SLASH} & \boxed{1} \end{bmatrix}, \boxed{4}\begin{bmatrix} \text{LOC} & \text{S}[\textit{fin}] \\ \text{SLASH} & \boxed{2} \end{bmatrix} \right\rangle \\ \text{BIND} & \{\,\} \\ \text{SLASH} & \boxed{1} \cup \boxed{2} \end{bmatrix}$$

The configurational passing of SLASH features can now be simplified to a constraint which only mentions the head daughter:

(22) **SLASH Inheritance**:

$$\textit{hd-val-ph} \Rightarrow \begin{bmatrix} \text{SLASH} & \boxed{1} \\ \text{HD-DTR} & \left[\text{SLASH } \boxed{1}\right] \end{bmatrix}$$

SLASH Inheritance is defined as a constraint on *head-val-phrases*, which is the type of phrases involving only head, complement, or subject daughters, but, crucially, no filler daughters (see Sag 1997).[9] For the latter type of phrase, we assume the following:[10]

[8] Following Sag (1997), we also eliminate the feature NONLOCAL that was introduced in PS-94.

[9] In Ginzburg and Sag (pear), types like *hd-val-ph* are eliminated and SLASH inheritance is a consequence of the Generalized Head Feature Principle, which identifies the entire SYNSEM value of head daughter and mother—by default.

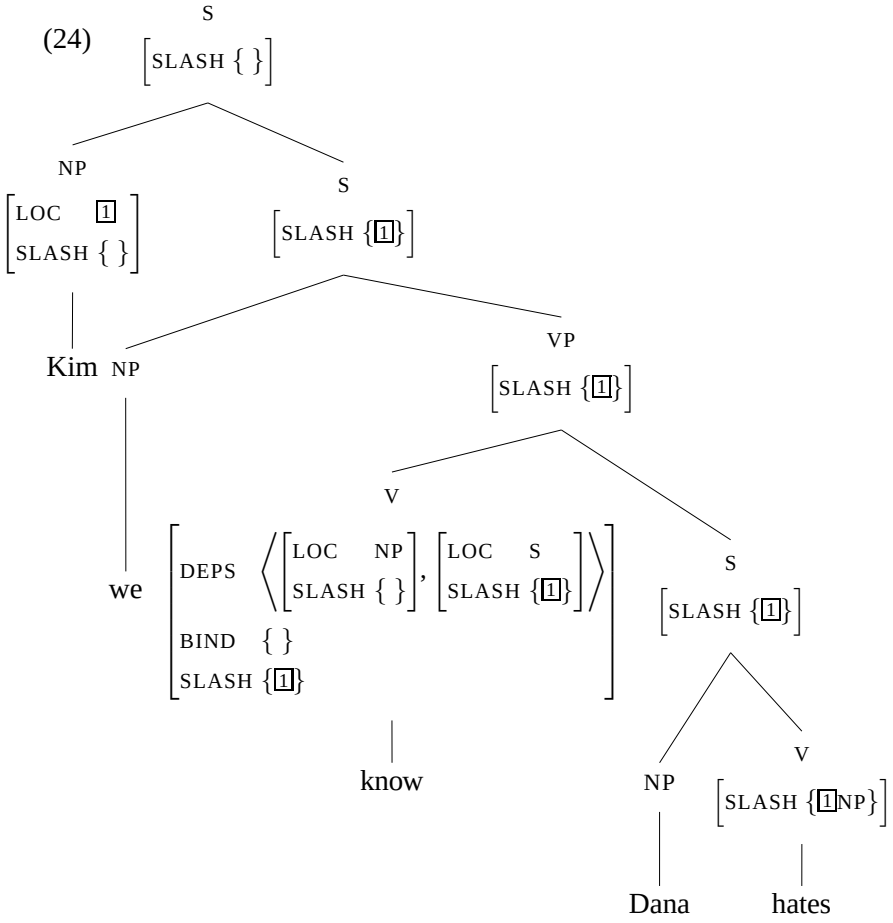
[10] Here '$\uplus$' designates the operation of disjoint set union, which is just like familiar set union except that the disjoint union of two sets with a nonempty intersection is undefined.

(23)

$$\textit{hd-filler-ph} \Rightarrow \begin{bmatrix} \text{SUBJ} & \langle\ \rangle \\ \text{SLASH} & \boxed{2} \uplus \boxed{3} \\ \text{HD-DTR} & \left[\text{SLASH } \boxed{2} \uplus \{\boxed{1}\}\right] \\ \text{NON-HD-DTRS} & \left\langle \begin{bmatrix} \text{LOC} & \boxed{1} \\ \text{SLASH} & \boxed{3} \end{bmatrix} \right\rangle \end{bmatrix}$$

The SLASH value of a head/filler phrase is the SLASH value of the head daughter minus the filler plus the SLASH value of the filler. Thus, we obtain structures such as (24).

(24)



What remains to be explained is how the bottom of the extraction path, e.g. the verb *hates* in (24), may introduce a non-empty value for SLASH. Traceless accounts of extraction within HPSG have posited a lexical rule to account for this. The Complement Extraction Lexical Rules of PS-94, Sag and Fodor (1994) or Sag (1997), for instance, all remove an element from

COMPS, and at the same time ensure (either directly or indirectly) that the value of SLASH on the output of the rule will include the local features of the removed complement. Such a rule is awkward, as it uses a device which is primarily intended to account for processes that are clearly morphological in nature. The usual diagnostic for distinguishing lexical from syntactic rules is that only the former have exceptions which are triggered by the lexical material involved. Many languages lexically mark words with a non-empty SLASH value, but to the best of our knowledge this is never the case for complement extraction per se. Furthermore, the formal status of the Complement Extraction Lexical Rule has itself been the matter of some debate (Hinrichs and Nakazawa, 1996; Höhle, 1995; Müller, 1996).

The key insight to our solution is that given a constraint-based lexical architecture in which ARG-ST lists are associated with valence features such as SUBJ and COMPS, there is no need for rules that remove elements from the valence lists; one can simply avoid putting them there in the first place. Thus extraction is located on the interface between the lexicon and the syntax rather than in either one of these components individually.

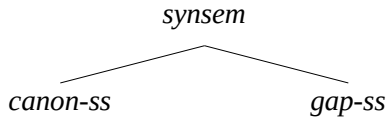
As illustrated in section 2, our analysis follows Sag (1997) in defining *gap-ss* as a new subtype of *synsem*. The LOCAL value of a *gap-ss* element corresponds to its SLASH value:

(25) $$\textit{gap-ss} \Rightarrow \begin{bmatrix} \text{LOC} & \boxed{1} \\ \text{SLASH} & \{\boxed{1}\} \end{bmatrix}$$

Elements of type *gap-ss* in this theory play a role similar to that of traces in a movement theory, with an important difference. Unlike traces, *gap-ss* elements exist only on the list of dependents of the lexical head which selects them. They play no independent role in the phrase structure, linear order, or phonology of phrasal signs.

The proper distinction between gaps and other, canonical, *synsem* types can be expressed in terms of the following type hierarchy:

(26)



The question of whether a dependent of a lexical head is to be realized as a local dependent (i.e. on one of the valence features SUBJ or COMPS) or a nonlocal dependent is a consequence of the Argument Realization constraint in (11), repeated in (27).

(27) **Argument Realization**:

$$\textit{word} \Rightarrow \begin{bmatrix} \text{SUBJ} & \boxed{1} \\ \text{COMPS} & \boxed{2} \ominus \textit{list(gap-ss)} \\ \text{DEPS} & \boxed{1} \oplus \boxed{2} \end{bmatrix}$$

Argument Realization, together with the principle of Canonicality in (28), ensures that none of the *gap-ss* elements on the DEPS list appears on the COMPS list.

(28) **Canonicality**:

$$sign \Rightarrow \left[\text{SYNSEM } canon\text{-}ss\right]$$

As lexemes will typically specify that their arguments are of type *synsem* and are not either of type *canon-ss* or of *gap-ss*, Argument Realization opens the possibility of realizing a non-subject dependent either as a complement or as a gap. For instance, the feature structure in (29) can satisfy Argument Realization either as shown in (30a) or as shown in (30b):

(29)
$$\begin{bmatrix} \text{I-FORM } \textit{hates} \\ \text{SUBJ} \quad \langle \textit{synsem} \rangle \\ \text{DEPS} \quad \left\langle \begin{bmatrix} \text{LOC} \quad \text{NP} \\ \text{SLASH } \boxed{1} \end{bmatrix}, \begin{bmatrix} \text{LOC} \quad \text{NP[acc]} \\ \text{SLASH } \boxed{2} \end{bmatrix} \right\rangle \\ \text{SLASH} \quad \boxed{1} \cup \boxed{2} \end{bmatrix}$$

(30) a.
$$\begin{bmatrix} \text{I-FORM } \textit{hates} \\ \text{SUBJ} \quad \left\langle \boxed{3} \begin{bmatrix} \text{LOC} \quad \text{NP} \\ \text{SLASH } \boxed{1} \end{bmatrix} \right\rangle \\ \text{COMPS} \quad \left\langle \boxed{4} \begin{bmatrix} \text{LOC} \quad \text{NP}\left[\textit{acc}\right] \\ \text{SLASH } \boxed{2} \end{bmatrix} \right\rangle \\ \text{DEPS} \quad \langle \boxed{3}, \boxed{4} \rangle \\ \text{SLASH} \quad \boxed{1} \cup \boxed{2} \end{bmatrix}$$

b.
$$\begin{bmatrix} \text{I-FORM } \textit{hates} \\ \text{SUBJ} \quad \left\langle \boxed{3} \begin{bmatrix} \text{LOC} \quad \text{NP} \\ \text{SLASH } \boxed{1} \end{bmatrix} \right\rangle \\ \text{COMPS} \quad \langle \ \rangle \\ \text{DEPS} \quad \left\langle \boxed{3}, \begin{bmatrix} \textit{gap-ss} \\ \text{LOC} \quad \boxed{2} \\ \text{SLASH } \{\boxed{2}\} \end{bmatrix} \right\rangle \\ \text{SLASH} \quad \boxed{1} \cup \left\{ \boxed{2} \text{NP}[\textit{acc}] \right\} \end{bmatrix}$$

In (30a), the complement is realized on COMPS. This is the kind of structure which occurs in sentences without extraction (e.g. (31a)), but also in

sentences where the extracted element is inside the object complement (as in (31b)). Note that because the SYNSEM value of any overt element (i.e. a sign) is of type *canon-ss* (by the principle of Canonicality in (28)), the complement *synsem* of (30a) must be of type *canon-ss* in all head-complement phrases headed by (30a). In (30b), the object dependent is realized as a gap and thus is extracted, as in (24) or (31c).

(31) a. She hates every inch of this school.

b. Which school does Dana say she hates [every inch of __]?

c. Which school does Kim think she hates __?

Even though lexemes typically do not specify their arguments as being of type *gap-ss* or *canon-ss*, this does not mean that there cannot be lexemes imposing more specific constraints on the type of their arguments. For instance, one of the usages of the verb *assure* requires an NP-object which cannot be realized as a local dependent (Kayne, 1980):

(32) a. This candidate, they assured me to be reliable.

b.*They assured me this candidate to be reliable.

We can account for the contrast in (32) by assuming that the direct object of *assure* (or simply the subject of the infinitival VP complement) is specified as being of type *gap-ss*.

Note that the identification of the LOCAL value of the gap with its SLASH value (and, ultimately, the LOCAL value of the filler) is simply stipulated by the constraint in (23). Nothing else about this analysis requires that the filler be identical to the gap in all respects. This provides a possible means to deal with examples in which the filler does not match the selectional restrictions imposed on the gap, as in examples like (33).

(33) a. You can rely on Chris.

b.*You can rely on that Chris will come.

c. Chris, you can rely on __.

d. That Chris will come, you can rely on __.

In fact, Webelhuth (1992) argues that in general sentential fillers correspond to NP gaps. One way that this generalization could be expressed under the analysis of extraction presented here is by positing a new type of gap with the constraint in (34):[11]

[11] A further constraint may be necessary if mere semantic selection (which we take to be extragrammatical in general) is insufficient to explain the deviance of examples like *#That Chris will come, you can own*.

(34) *sgap-ss* $\Rightarrow$ $\begin{bmatrix} \text{LOC} & \text{NP} \\ \text{SLASH} & \{\text{S}\} \end{bmatrix}$

So, the additional flexibility introduced by the constraints on gaps allows us to account for examples of limited connectivity between fillers and gaps that provide a serious challenge to standard movement-based treatments of extraction.[12]

3.2. CROSS-LINGUISTIC SUPPORT

As a further illustration of our approach, consider how it might account for one of the phenomena mentioned in the introduction. As McCloskey (1979, 1989) notes, Irish has two different complementizer particles, *goN* and *aL*. The only difference between the two is that the former cannot appear in a clause out of which something has been extracted, whereas the latter can *only* appear in a clause out of which something has been extracted. The pattern is illustrated in (35).

(35) a. Shíl mé *goN* mbeadh sé ann
thought I PRT would-be he there
'I thought that he would be there.'

b. Dúirt mé *gurL* shíl mé *goN* mbeadh sé ann
said I goN+PAST thought I PRT would-be he there
'I said that I thought that he would be there.'

c. an fear *aL* shíl mé *aL* bheadh _ ann
[the man]$_j$ PRT thought I PRT would-be _$_j$ there
'the man that I thought would be there'

d. an fear *aL* dúirt mé *aL* shíl mé *aL* bheadh _ ann
[the man]$_j$ PRT said I PRT thought I PRT would-be _$_j$ there
'The man that I said I thought would be there'

e. an fear *aL* shíl _ *goN* mbeadh sé ann
[the man]$_j$ PRT thought _$_j$ PRT would-be he there
'the man that thought he would be there'

f. an fear *aL* dúirt sé *aL* shíl _ *goN* mbeadh sé ann
[the man]$_j$ PRT said he PRT thought _$_j$ PRT would-be he there
'the man that he said thought he would be there'

[12] See Bresnan (2000) for more discussion of mismatches of this sort.

Sells (1984) argues that the putative complementizer particles *aL* and *goN* are pre-verbal particles and not sentential complementizers, citing both morphosyntactic and syntactic evidence for this analysis. For example, verbal constituents including the particle can be coordinated, whereas coordination of the bare verbs is impossible:

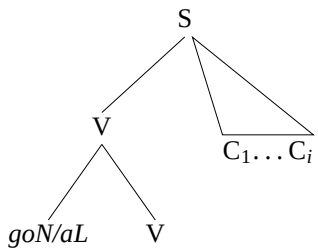
(36) a. an fear *aL* cheannaionn agus *aL* dhíolann tithe
the man PRT buys and PRT sells houses

the man that buys and sells houses

b.*an fear *aL* cheannaionn agus dhíolann tithe

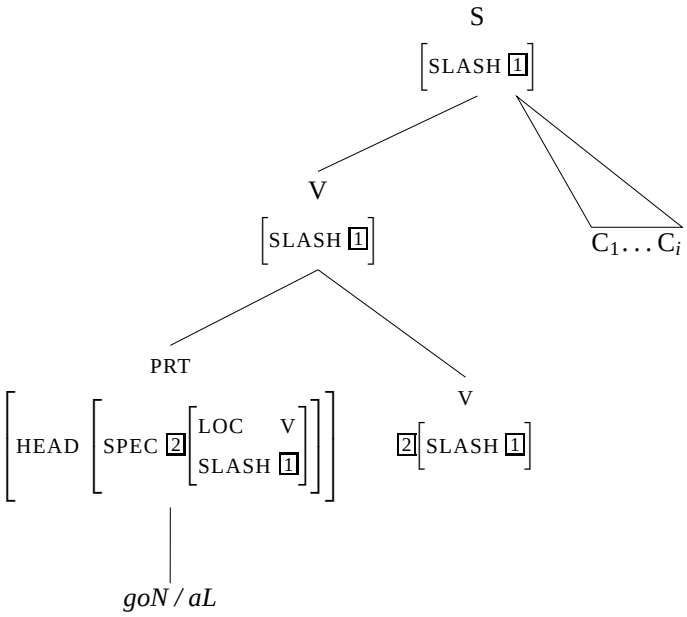
Thus, Sells proposes the following structure:

(37)



Within HPSG, these pre-verbal complementizer particles can be analyzed as markers which select for a lexical verb via the SPEC feature:

(38)



The choice of marker in (38) depends on whether 1 is empty or not. This can be accounted for simply by assuming that *aL* selects for a verb with a

nonempty SLASH value, whereas *goN* selects a verb with an empty SLASH value:[13]

(39) a.
$$\begin{bmatrix} \text{ROOT} & \textit{goN} \\ \text{HEAD} & \begin{bmatrix} \textit{prt} \\ \text{SPEC} & \begin{bmatrix} \text{LOC} & \text{V} \\ \text{SLASH} & \textit{eset} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

b.
$$\begin{bmatrix} \text{ROOT} & \textit{aL} \\ \text{HEAD} & \begin{bmatrix} \textit{prt} \\ \text{SPEC} & \begin{bmatrix} \text{LOC} & \text{V} \\ \text{SLASH} & \textit{neset} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

The simplicity of this account is a direct consequence of the fact that configurational passing of nonlocal features is replaced by a lexicalist SLASH Amalgamation constraint, which ensures that a verb has a non-empty SLASH value just in case one of its complements is slashed. In an approach using the Nonlocal or Foot Feature Principle, on the other hand, the SLASH value of the verb would be the empty set in both cases. Consequently, the distinction between gapped and complete clauses could only be made if complementizer particles could select for the SLASH values of the elements on the COMPS list of the verbs they combine with. This highly marked nonlocal selection criterion is avoided in the lexicalist account.[14]

Similarly, the type of Chamorro morphology illustrated in (4), and repeated in (40), registers the fact that a verb's subject either is extracted or is realized but contains a gap.

(40) a. Hayi *f-um-a'gasi* i kareta
who WH.SU-wash the car

'Who washed the car?'

b. Hayi si Juan ha-sangan-i hao [*f-um-a'gasi* i kareta]
who UNM Juan E3S-say-DAT you WH.SU-wash the car

'Who did Juan tell you washed the car?'

[13] We use *eset* and *neset* for the types of the empty and nonempty set, respectively.

[14] As an anonymous reviewer points out, this analysis is also much simpler than approaches proposed within the Minimalist Program (e.g. Collins, 1997) which require two separate features, one to trigger movement and one to trigger agreement.

c. Hafa *um-istotba* hao [ni malagao'-na i
what WH.SU-disturb you COMP WH.OBL-want-3SG the
lahi-mu]
son-your

'What does it disturb you that your son wants?'

We may treat such verbs as simply requiring that their most prominent dependent be slashed:

(41) a. $\begin{bmatrix} \text{I-FORM } \textit{f-um-a'gasi} \\ \text{DEPS} \quad \left\langle \begin{bmatrix} \text{LOC} \quad \text{NP} \\ \text{SLASH } \textit{neset} \end{bmatrix}, \text{NP} \right\rangle \end{bmatrix}$

b. $\begin{bmatrix} \text{I-FORM } \textit{um-istotba} \\ \text{DEPS} \quad \left\langle \begin{bmatrix} \text{LOC} \quad \text{S} \\ \text{SLASH } \textit{neset} \end{bmatrix}, \text{NP} \right\rangle \end{bmatrix}$

3.3. TOUGH GUYS

The head-driven approach to SLASH amalgamation has the further advantage that it accounts for the fact that the binding of SLASH is not restricted to head-filler constructions, but sometimes is triggered by lexical items, even if the SLASH value which gets bound originates arbitrarily deep within a dependent of the binder. We discuss two of these cases below.

Adjectives such as *easy* or *tough* in (42) are standardly analyzed as selecting for an infinitival complement missing an NP.

(42) a. Kim is easy to please __.

b. That journal is tough to provide camera-ready material for __.

Following PS-94 we may account for this phenomenon by assuming that the feature (TO-)BIND takes a nonempty value for adjectives like *easy*.

(43) $\begin{bmatrix} \text{ROOT} \quad \textit{easy} \\ \text{ARG-ST} \left\langle \text{XP}_i, \begin{bmatrix} \text{LOC} \quad \text{CP}[\textit{to}] \\ \text{SLASH } \{\boxed{1}\text{NP}[\textit{acc}]_i\} \uplus \boxed{2} \end{bmatrix} \right\rangle \\ \text{BIND} \quad \{\boxed{1}\} \end{bmatrix}$

The BIND feature will ensure that the NP[*acc*] gap will not be amalgamated into the SLASH value of *easy* itself.

Note, however, that other SLASH values within the VP are amalgamated as usual, thus accounting for cases such as (44), where the phrase *easy to play on* has a nonempty SLASH value.

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

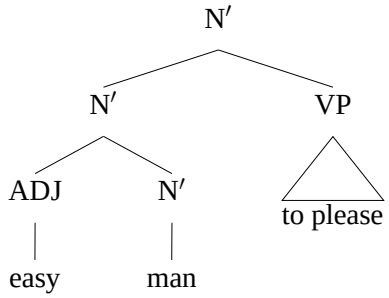
3.4. SUBBINDING

The analysis just presented also provides an account of more complicated cases, such as those noted in Flickinger and Nerbonne (1992):

(45) an easy man to please __

Flickinger and Nerbonne argue for the following constituent structure:

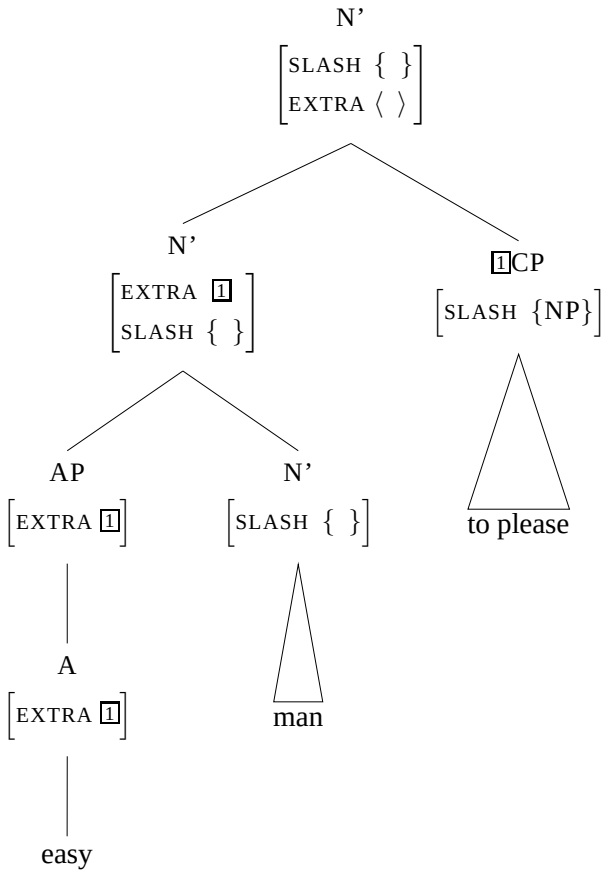
(46)



The adjective *easy* is not the head of this structure, nor is it the head of the substructure *easy man*. A configurational approach to SLASH amalgamation makes the incorrect prediction that the higher N′ in (46) should be slashed, as one of its daughters (*to please*) is slashed. That is, the BIND feature on *easy*, which appears in a subordinate structure, cannot be used to ensure that the phrase as a whole must bear an empty SLASH specification. The lexicalized SLASH amalgamation approach, by contrast, faces no such problem. The SLASH value of the VP is bound lexically by the adjective *easy*. The fact that the phrase *to please* is slashed in (46) does not require that the same be true of the dominating N′, whose SLASH value is identified with that of its head daughter. Of course, this leaves an important puzzle concerning the construction in (46) unexplained, namely the fact that a complement of *easy* appears in a superordinate syntactic position. Flickinger and Nerbonne suggest a principle of 'subcategorization transfer' to account for this, which allows *to please* also to function as a complement of the noun *man*. We suspect an account in terms of a more general treatment of extraposition is preferable. Such an analysis, examination of which would take us too far afield, might constrain the distribution of the EXTRA feature as follows:[15]

[15] This approach presupposes that the lexical entry for the *attributive* adjective *easy* is as sketched in (i).

(47)



Another puzzling construction is *en*-clitization in French. Miller and Sag (1997) argue for a theory of cliticization in which arguments may be realized either as selected-for dependents or as clitics of the verb on whose argument structure they appear. Clitic climbing is accounted for by an appeal to argument composition: verbs which allow clitics corresponding to arguments of an embedded verb are actually argument composition verbs whose ARG-ST list includes elements inherited from the ARG-ST list of their verbal complement. However, *en*-clitization, illustrated in (48) below, is different.

(i) *easy*: $\begin{bmatrix} \text{BIND} & \{\boxed{2}\} \\ \text{SLASH} & \{\,\} \\ \text{EXTRA} & \left\langle \left(\begin{bmatrix} \text{LOC} & \text{CP}[\textit{to}] \\ \text{SLASH} & \{\boxed{2}\} \end{bmatrix} \right) \right\rangle \end{bmatrix}$

(48) a. Marie en connaît la fin.
Marie of-it knows the end

'Marie knows the end of it.'

b. La fin en est désagréable.
the end of-it is unpleasant

'The end of it is unpleasant.'

The clitic *en* corresponds to a *de*-phrase which is a semantic argument of a noun, and the phenomenon is far more restricted than other types of clitization (it is restricted to a formal register and occurs only with arguments of the right semantic type). On the other hand, the construction does share properties with extraction of *de*-phrases:

(49) a. J'ai écouté le garçon avec la clef.
I have heard the boy with the key

'I heard the boy with the key.'

b. *la clef$_j$, avec laquelle$_j$ j'ai écouté le garçon...
the key with which I have heard the boy...

'the key with which I heard the boy...'

(50) a. Marie lit la fin du livre.
Marie reads the end of-the book

'Marie is reading the end of the book.'

b. le livre dont Marie lit la fin...
the book of-which Marie reads the end...

'the book of which Marie is reading the end...'

c. La fin du livre est désagréable.
the end of-the book is unpleasant

'The end of the book is unpleasant.'

d. le livre dont la fin est désagréable...
the book of-which the end is unpleasant...

'the book of which the end is unpleasant...'

Extraction out of NPs in general is excluded in French, as in (49), the only exception being extraction of certain *de*-phrases.

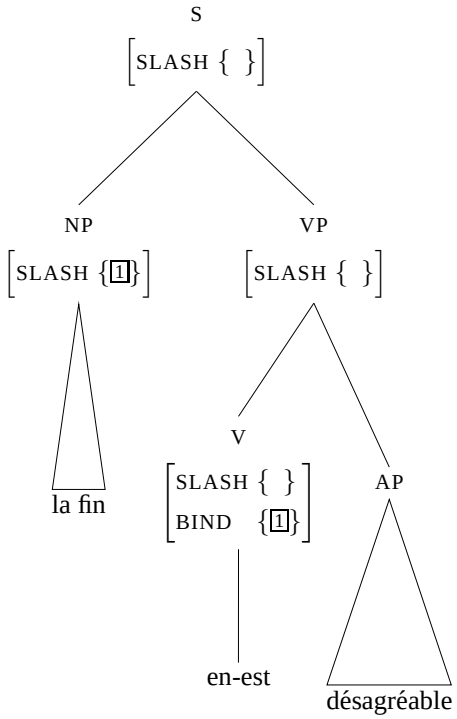
The similarity between *en*-cliticization and extraction of *de*-phrases leads Miller and Sag to propose a lexical rule which derives *en*-cliticized verbs. The

output of this rule is a verb which selects for a subject whose SLASH value contains a *de*-phrase. Furthermore, the verb lexically binds the SLASH value of the subject. For instance the en-cliticized form of copular *être* would be:

(51) *en-est*: $\begin{bmatrix} \text{SUBJ} & \left\langle \left[\text{SLASH} \{ \boxed{1} \text{NP}[de] \} \right] \right\rangle \\ \text{COMPS} & \langle \text{AP} \rangle \\ \text{BIND} & \{ \boxed{1} \} \end{bmatrix}$

Note that this entry binds the SLASH value of a dependent (e.g. the subject) which is typically introduced higher in the phrase structure than the verb, as sketched in (52).

(52)



This sort of structure is compatible with our head-driven account of SLASH-amalgmation, but appears to be extremely problematic for previous theories, where the SLASH binder must always be in a syntactically superior position to the extraction site.[16] Indeed, these facts stand as a challenge to all theories whose constraints require binders to be structurally superior to (e.g. c-commanding) the traces that they bind.

[16] But see Chae (1992).

3.5. WITHOUT A TRACE

Phonetically empty elements are a familiar ingredient of syntactic analyses of extraction. However, at least within nontransformational frameworks, there is also a tradition of traceless approaches to unbounded dependency constructions (Gazdar et al., 1984; Steedman, 1996; Morrill, 1994; Kaplan and Zaenen, 1989). In Sag (1998) (see also Sag and Fodor (1994)), the most compelling arguments in favor of traces which have been proposed over the years are reviewed. These include so-called *wanna* contraction, auxiliary contraction, and both strong and weak crossover phenomena. Drawing on a number of recent results, the conclusion reached by Sag is that all of the existing arguments for traces face serious problems. Moreover, there is further evidence that appears to be more readily compatible with a theory of extraction which does not rely on traces. We review this latter evidence here briefly.

3.5.1. *Floated Quantifiers*

Floated quantifiers may not appear directly before an extraction site:

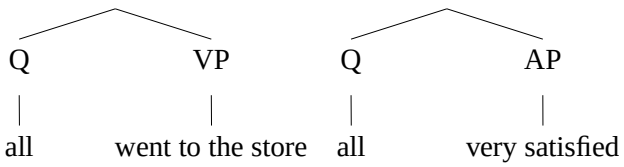
(53) a. They (all) were (all) completely satisfied.

b. How satisfied do you think they all were _?

c.*How satisfied do you think they were all _?

While certain accounts of these examples have involved stipulations making reference to traces, the alternative account suggested by Sag and Fodor (1994) does not. It is based on the analysis of floated quantifiers developed by Dowty and Brodie (1984) in which such quantifiers, like certain adverbial modifiers, are base-generated as VP or AP-adjoined modifiers:[17]

(54) a.



b.

Brodie and Dowty's account is consistent with a range of semantic constraints on quantifier floating, and also appears to provide a direct explanation for the problematic cases above—given a traceless account of extraction. On a traceless extraction analysis, there is no way to generate a sentence like (53c), as there is no empty constituent for the quantifier to attach to.

[17] See (Bobaljik, 1998) for a recent review. Sportiche (1988) objects to the dual categorization of quantifiers he sees in Dowty and Brodie's account, yet—by his own admission (p. 444)— can provide no account of facts like (53c).

3.5.2. *The Conjunct Constraint*

Sag (1998) shows that the assumption that there are no *wh*-traces provides an immediate explanation for the ungrammaticality of examples like (55a–c), which are violations of the Conjunct Constraint, part of the Coordinate Structure Constraint first formulated by Ross (1967).[18]

(55) a.*Which of her books did you find both [[a review of __] and [__]]?

b.*Which of her books did you find [[__] and [a review of __]]?

c.*Which rock legend would it be ridiculous to compare [[__] and [__]]?
(cf. Which rock legend would it be ridiculous to compare __ with himself?)

Sag's reasoning is as follows:

– A *wh*-gap is simply a syntactic position where a dependent of a head fails to be realized (rather than a position where a phonetically empty constituent is syntactically realized).

– Coordinate structures are unheaded and the elements that are coordinated, i.e. the conjuncts of a coordinate structure, must be syntactic constituents (or perhaps sequences thereof).

– Therefore, it follows that *wh*-gaps, which are not constituents, can never be conjuncts.

This result is achieved without stipulation in ANY traceless analysis of extraction phenomena, including the one developed here, as long as the conjunction is not treated as the head of the coordinate phrase.[19]

3.5.3. *Processing Complexity*

Pickering and Barry (1991) argue on psycholinguistic grounds for a theory of extraction in which an extracted element is associated with its semantic

[18] Goodall (1987) seeks to explain these examples as Principle C violations, but the severe ungrammaticality of the examples in (55) makes such an account quite unlikely. On Goodall's approach, the examples in (55) should be no worse than sentences like *We invited Jack's$_i$ friends and him$_i$ both*, which seem perfectly acceptable.

[19] The suggestion that conjunctions are heads, which reappears from time to time in the literature (see, for example, Rothstein (1991) and Munn (1992)), is not particularly intuitive, as the category of the coordinate phrase (as reflected by its outward distribution) is determined by that of the conjuncts, rather than the conjunction. Hence the 'conjunction-as-head' analysis requires an unprecedented chameleon-like categorial behavior on the part of the putative head. There are further objections that could be made against this analysis (for example, the difficulty in accounting for complex coordinations with *both*/*and*, *neither*/*nor*, and the like). We will not pursue these matters here, however, adopting instead the traditional view that coordinate structures instantiate a *sui generis* construction type.

role as soon as the head licensing this element is encountered. Often, the extraction site and the head licensing it are adjacent, as in (56a). This is not the case for (56b).

(56) a. The policeman saw the boy that the crowd at the party accused __ of the crime.

b. That's the prize which we gave [every student capable of answering every single tricky question on the details of the new and extremely complicated theory about the causes of political instability in small nations with a history of military rulers] __.

Surprisingly, such examples are relatively easy to process, unlike examples in which a constituent follows an extremely long preceding constituent, as in (57).

(57) We gave [every student capable of answering every single tricky question on the details of the new and extremely complicated theory about the causes of political instability in small nations with a history of military rulers] [a prize].

If interpretation of extracted elements correlates with the position of a trace at the extraction site, this contrast remains unexplained.

But if interpretation can take place when the head selecting for the extracted element is encountered, then this should not be surprising. For example, in our traceless analysis of filler-gap constituents, it is precisely when the verbal head is encountered that the 'extracted' element in working memory can be identified with an argument structure position. It is the verb that terminates the processing of the extraction dependency, not a trace position that might occur several words after the verb.

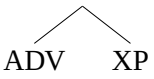
It should be noted that there is more recent work (Gorrell, 1993; Gibson and Hickok, 1993) that tries to reconcile such facts with a trace-based analysis by making specific (head-driven) assumptions about the structures posited by the human sentence processor. The fact remains, however, that the predictions of any such approach are at best just those that follow immediately from the traceless analysis of extraction dependencies.

In sum: (1) There is no independent evidence supporting the existence of *wh*-trace (as argued by Sag (1998) and Sag and Fodor (1994)). (2) There are further phenomena—those just surveyed—that are immediately explained from the assumption that there are no *wh*-traces. (3) Therefore, the traceless nature of our extraction analysis provides further support for its correctness.

4. The Analysis of Adjuncts

The idea that adjuncts combine syntactically with the phrases that they modify semantically and form modifier-head structures like (58) is firmly entrenched in modern grammatical theory.[20]

(58)

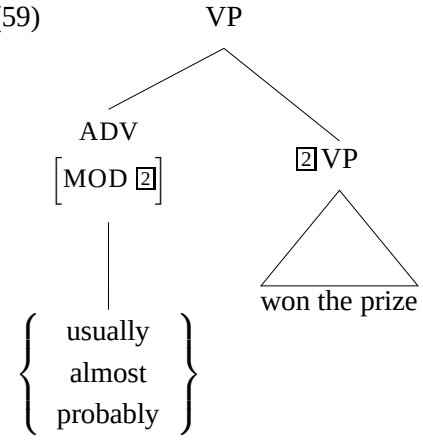


The semantics that goes with such phrases generally involves the adverbial's content taking that of the phrase it modifies as its semantic argument.

In HPSG terms, adjuncts are endowed with a specification for the feature MOD, whose value must be identified with (the SYNSEM of) the head daughter in a Head-Modifier Structure like (58). In this way, the fact that adjectives modify nominal expressions, adverbials modify verbal expressions, etc. can all be reduced to lexical specifications associated with particular kinds of modifiers.

And indeed we have no reason to question this traditional wisdom in the case of preverbal adverbs, which we will assume occur as modifiers in structures like (59), given the lexical entry in (60).

(59)



[20] These are often referred to as 'head-adjunct' structures, a terminology we will avoid here in order to distinguish between elements that belong to the DEPS list ('adjuncts') and those that appear in modifier-head structures ('syntactic modifiers'). Excluded from discussion here are all sorts of elements often called modifiers, e.g., parentheticals, extraposed phrases, result clauses, and so forth.

(60)
$$\begin{bmatrix} \text{I-FORM } \textit{usually} & \\ \text{HEAD} & \begin{bmatrix} \textit{adv} & \\ \text{MOD} & \begin{bmatrix} \text{HEAD V} \\ \text{CONT}\,|\,\text{KEY } \boxed{1} \end{bmatrix} \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{KEY} & \begin{bmatrix} \textit{usually\_rel} \\ \text{ARG } \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

This type of analysis is adequate for a wide range of cases to which it is standardly applied. However, we find in many languages types of adverbials that defy any simple analysis in terms of the syntactic combination of modifiers and heads. In particular, it has been argued that cliticization (Miller, 1992), word order (van Noord and Bouma, 1994; Abeillé and Godard, 1997), scope (Manning et al., 1999; van Noord and Bouma, 1994; Kim and Sag, 1995; Przepiórkowski, 1999a), and case marking (Przepiórkowski, 1999b; Przepiórkowski, 1999c) suggest that certain adverbial phrases must be selected for by the same mechanism which accounts for the selection of complements.[21]

For example, it has been argued (Manning et al., 1999; van Noord and Bouma, 1994) that a treatment of adjuncts as complements provides an account for scope ambiguities that may be observed in constructions involving complex predicates. Consider, for instance, the following examples:

(61) a. . . . dat Marie Jan dikwijls een boek liet lezen
that Mary John often a book made read

'. . . that Mary (often) made John (often) read a book' (Dutch)

b. often(cause(mary,read(john,book)))

c. cause(mary,often(read(john,book)))

(62) a. Noriko-ga Masaru-ni gakkoo-de hasir-ase-ta
Noriko-NOM Masaru-ACC school-at run-CAUS-PAST

'Noriko made Masaru run at school.' (Japanese)

b. at-school(cause(noriko,run(masaru)))

c. cause(noriko,at-school(run(masaru)))

[21] The treatment of adverbials as syntactic dependents dates back at least to the medieval Arabic grammarians (see, e.g., Bohas, 1990). Similar proposals have occasionally made their way into the linguistic literature at large; see, for example, Tesnière (1959), Melčuk (1979), and McConnell-Ginet (1982).

In each case, an adjunct may either take wide scope over a complex predicate or else narrow scope over only a part of this predicate. While the wide scope readings are easily accounted for in practically any treatment of adjuncts, the narrow scope readings are not. An analysis of the verb cluster *liet lezen* in (61) or the complex verb *hasiraseta* in (62) involving incorporation, reanalysis, or argument composition would fail to provide an appropriate syntactic subconstituent for the adverb to modify to yield the narrow scope reading. On the other hand, if adjuncts are on the COMPS list of the verb they modify, the ambiguity arises because the adverbial could be on the COMPS list of (and take scope over) either the governing or the embedded verb. Furthermore, composition of lists of complements, as proposed on entirely independent grounds by Hinrichs and Nakazawa (1994), accounts for the fact that both narrow and wide scope adverbials may precede complex predicates like the Japanese causative construction and the Dutch verb cluster.[22]

Another challenge to the standard view of adverbial modifiers comes from the behavior of case marking in languages like Finnish, Korean, and Polish (Maling, 1989; Maling, 1993; Wechsler and Lee, 1996; Przepiórkowski, 1999b; Przepiórkowski, 1999c). For example, consider the following Finnish sentences:

(63) a. Liisa muisti matkan vuoden
Liisa.NOM remembered trip.ACC year.ACC

'Liisa remembered the trip for a year.'

b. Lapsen täytyy lukea kirja kolmannen kerran
child.GEN must read book.NOM third time.ACC

'The child must read the book for a third time.'

c. Kekkoseen luotettiin yksi kerta
Kekkonen.ILL trust.PASS one time.NOM

'Kekkonen was trusted once.'

In Finnish, the least oblique argument of a verb is assigned nominative case, with other arguments receiving accusative case, as in (63a). If a verb assigns quirky case to its least oblique argument, the next highest dependent gets the nominative case, as in (63b). This is true even if, as in (63c), the next available dependent happens to be an adverbial. It is difficult to see how this generalization can be adequately expressed if Finnish verbs do not select for at least some kinds of adverbial modifiers. There must be some analogue of our DEPS list which imposes a uniform, hierarchical structure

[22] Under the analysis proposed by Manning et al. (1999), the narrow-scope adjuncts are actually on the ARG-ST lists of both the verbal stem and the morphologically derived causative word.

on complements and the relevant adjuncts. For a more detailed discussion of the commonality of arguments and modifiers in these and a variety of other languages, including French (cliticization, word order, and scope), German (word order and scope ambiguities), Polish and Russian (case assignment), see Przepiórkowski (1999b; 1999c).

Accounts within HPSG to date have assumed a lexical rule such as (9b) (repeated in (64)) to account for the fact that the COMPS list of a verb must be allowed to contain an arbitrary number of adjuncts after the non-subject arguments:

(64) $$\begin{bmatrix} \textit{verb} \\ \text{SUBJ} \quad \boxed{1} \\ \text{COMPS} \quad \boxed{2} \\ \text{ARG-ST} \quad \boxed{1} \oplus \boxed{2} \end{bmatrix} \mapsto \begin{bmatrix} \text{SUBJ} \quad \boxed{1} \\ \text{COMPS} \quad \boxed{2} \oplus \langle \text{ 'adverbial' } \rangle \\ \text{ARG-ST} \quad \boxed{1} \oplus \boxed{2} \end{bmatrix}$$

In the constraint-based approach outlined in section 2, this rule is superfluous, as the association between argument structure and dependents already accounts for the fact that the DEPS list, and thus COMPS, may contain adverbial *synsem*s. The relevant argument structure extension constraint (a revision of (10) above) is given in (65).

(65) **Argument Structure Extension**:

$$\textit{verb} \Rightarrow \begin{bmatrix} \text{HEAD} \quad \boxed{3} \\ \text{DEPS} \quad \boxed{1} \oplus \textit{list}\left(\begin{bmatrix} \text{MOD} \begin{bmatrix} \text{HEAD} \quad \boxed{3} \\ \text{KEY} \quad \boxed{2} \end{bmatrix} \end{bmatrix}\right) \\ \text{ARG-ST} \quad \boxed{1} \\ \text{CONT} \mid \text{KEY} \quad \boxed{2} \end{bmatrix}$$

This allows a verb's DEPS list to contain any number of adverbials in addition to the verb's arguments. Moreover, the MOD|HEAD value of the adverbial is identified with the HEAD value of the verb on whose DEPS list the adverbial appears. We have imposed minimal restrictions on the adverbial elements that can be included in the verb's DEPS list, allowing adverbs, PPs, and the like. It is likely that further constraints should be imposed on such dependents, but here we leave their exact nature open. The semantics can be accounted for by simply unifying the adverbial's MOD|KEY value with the KEY relation introduced by the verb. Since scopal adverbials identify the key of their MOD value with their semantic argument (as in (60) above), this has the effect of making the verb the semantic argument of the adverbial.[23]

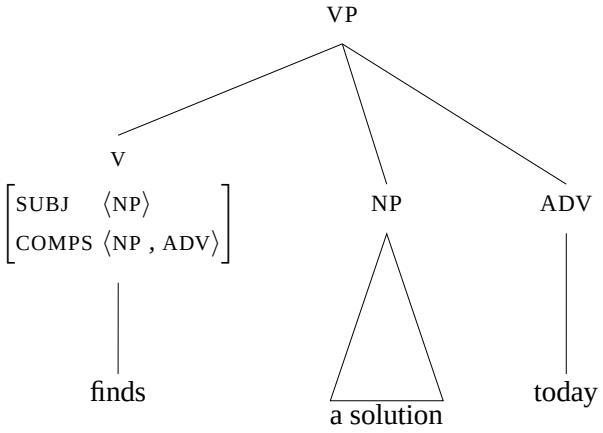
[23] In fact, a proper account of scope ambiguity requires that the semantics of the verb only be constrained to be *embedded within* the semantic structure identified as the argument of the adverbial phrase. See Bouma et al. (1998) for a detailed treatment of the semantics and scope of adverbial phrases in terms of Minimal Recursion Semantics (Copestake et al., 1999).

The combined effect of Argument Structure Extension and the Argument Realization constraint introduced earlier is that adverbials may appear as complements of the verb. For example, if a transitive verb such as *find* has an argument structure consisting of two NPs, the following is consistent with both Argument Structure Extension and Argument Realization:

(66)
$$\begin{bmatrix} \textit{trans} \,\&\, \textit{3sg} & \\ \text{I-FORM} & \textit{finds} \\ \text{HEAD} & \boxed{4}\textit{v}[\textit{fin}] \\ \text{SUBJ} & \langle\, \boxed{1}\text{NP}_{[3\text{sg}]} \,\rangle \\ \text{COMPS} & \left\langle \boxed{2}\text{NP} \,,\, \boxed{3}\begin{bmatrix} \text{MOD} & \begin{bmatrix} \text{HEAD} & \boxed{4} \\ \text{KEY} & \boxed{5} \end{bmatrix} \end{bmatrix} \right\rangle \\ \text{DEPS} & \langle\, \boxed{1} \,,\, \boxed{2} \,,\, \boxed{3} \,\rangle \\ \text{ARG-ST} & \langle\, \boxed{1} \,,\, \boxed{2} \,\rangle \\ \text{CONT}\,|\,\text{KEY} & \boxed{5} \end{bmatrix}$$

This lexical specification then licenses head-complement structures such as the following:

(67)



It is important to realize that we are not proposing to treat all adverbials as complements. It is generally the case that complements in English never precede the verb which locally selects for them. This suggests that the adverbs in the examples below cannot be complements.

(68) a. I think Kim almost found the solution.

b. Kim claimed that Sandy never sang for her.

Preverbal modifiers are then adjuncts which combine with VPs to give rise to modifier-head structures, as in (59). Postverbal adverbial phrases, on the

other hand, are complements and appear as non-head daughters in head-complement structures, as shown (67).[24]

We are also not proposing to eliminate the distinction between arguments and adjuncts. Arguments appear on ARG-ST, whereas adjuncts may only appear on DEPS. Thus Principle C of the binding theory outlined in PS-94 could remain exactly as formulated there, with the *o-command* relation defined in terms of the ARG-ST list, not the DEPS list. This allows o-command to distinguish between the argument PP in (69a) and the adverbial in (69b) for purposes of binding constraints.

(69) a.*I told them$_i$ about [the twins']$_i$ birthday.

b. I only get them$_i$ presents on [the twins']$_i$ birthday.

But in fact, there is mounting evidence showing that Principle C is more pragmatic in nature—as argued by Bolinger, Reinhart, and others (see Bresnan (2000) for an overview). Hence we would favor a binding theory that includes at most two principles: Principle A requiring a locally o-commanded anaphor to be locally o-bound and Principle B requiring a pronominal to be locally o-free. Examples like (70), cited by Hukari and Levine (1996b, 1996a) as evidence for a notion of command that includes both obliqueness and structural notions, are in our view not to be ruled out by binding theory.

(70)*He$_i$ always gets angry when Sandy$_i$ is criticized.

Rather, such examples violate only theme/rheme conditions on the anaphoric use of nonpronominal NPs. This is why examples like (71), whose grammatical structure is nondistinct from (70) (along relevant parameters) but whose information structure is radically different, permit the indicated coreference.

(71) He$_i$ gets angry whenEVER the people Sandy$_i$ loves criticize him.

Hence, any attempt to rule out examples like (70) via a version of Principle C that is based on grammatical structure (whether this is taken to be argument structure, phrase structure, or some combination of the two) will incorrectly rule out examples like (71). For this reason, we see no need to revise the notion of (local) o-command introduced by PS-94 along the lines suggested by Hukari and Levine.

[24] It is also worth pointing out that the configurational dichotomy between preverbal and postverbal adjuncts is consistent with observations concerning the relative scope of preverbal and postverbal adjuncts (see Bouma et al. (1998)).

5. Adjunct Extraction

There is considerable evidence, as argued in PS-94 and extensively in Hukari and Levine (1995), that adjuncts may be involved in unbounded dependency constructions. Hukari and Levine point out, for instance, that adjuncts not only occur in topicalizations and *wh*-questions, as in (72), but also in all other constructions involving unbounded dependencies, such as indirect questions, relatives, and clefts, as in (73).[25]

(72) a. On Tuesday, Sandy visits Leslie.

b. On Tuesday, I think it's likely that Sandy visits Leslie.

c. How often do you think Robin sees Kim?

(73) a. Kim wondered how she could repair the sink.

b. This is the restaurant in which Kim and Sandy first ordered couscous.

c. I will have lunch in whichever restaurant Leslie wants to have lunch.

d. It was in early January that Kim and Sandy first ordered couscous in a middle eastern restaurant.

Hukari and Levine (1995) further demonstrate that there is strong cross-linguistic evidence that the initial adjuncts in examples like (72) and (73) are part of an unbounded dependency construction (see also PS-94). Their argument rests on the observation that in all those cases where a language registers unbounded dependencies in some way, the registration occurs consistently with both arguments and adjuncts. This is shown for Irish below:

(74) a. Cén uair *aL* tháinig siad ńa bhaile _
which time$_j$ PRT came they home _$_j$

'what time did they come home'

b. Cén fáth *arL* dhúirt tú *aL* tháinig sé _
which reason$_j$ PRT said you COMP came he _$_j$

'why did you say he came'

[25] This is not to say, however, that *all* sententence-initial adverbials involve an unbounded dependency construction; see, for example, Foley and Van Valin (1984).

These examples show that when there is a preposed adjunct interpreted within the lower clause, the complementizer particle *aL* appears, signalling the presence of an unbounded dependency. Such observations indicate that attempts to account for apparent cases of adjunct extraction by assuming that they are in fact 'base-generated' and linked to their 'extraction site' by means of a semantic interpretation rule, must be misguided.

Hukari and Levine (1995) and PS-94 also shed considerable doubt on the common assumption that extraction out of adjuncts is possible only in so-called 'parasitic gap' constructions by presenting examples such as the following:

(75) a. What do you think Robin computed the answer with?

b. Who does Robin claim that Sandy sang a song for?

c. Which students is Roger capable of working totally independently of?

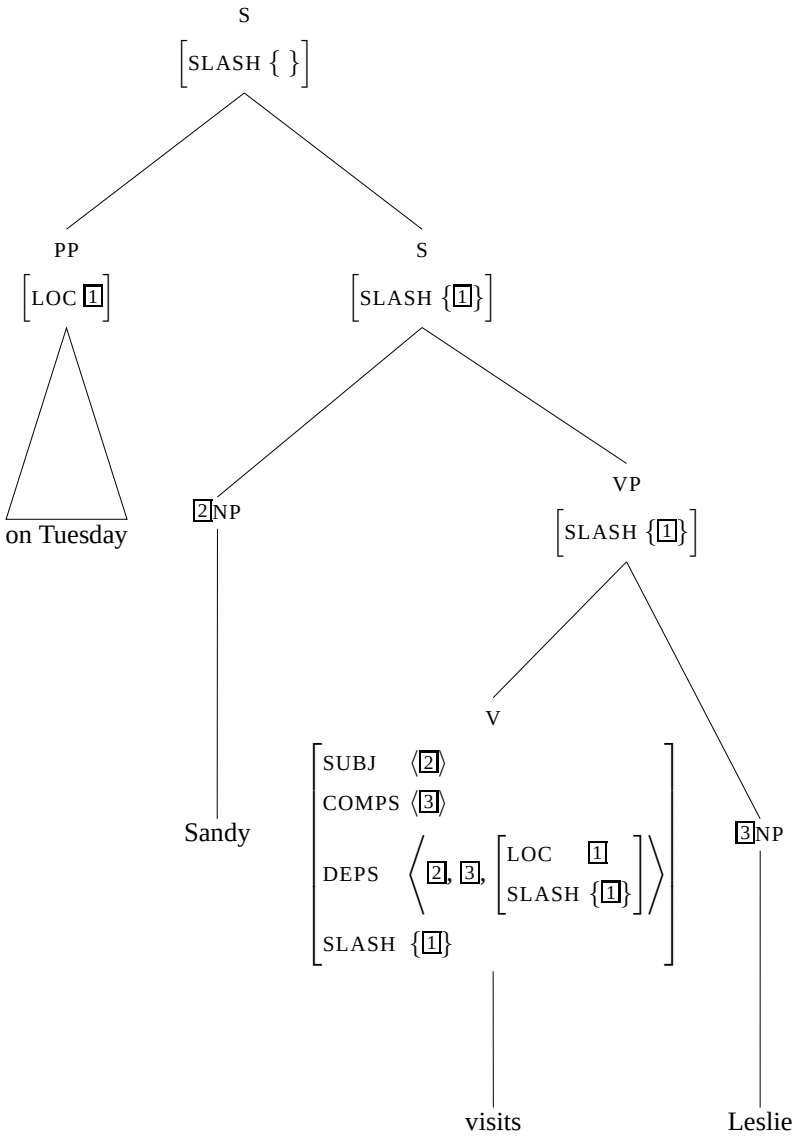
d. Which people can Robin run nearly as fast as?

e. Who does Kim write letters more frequently than?

The possibility of adjunct extraction and extraction out of adjuncts follow from the mechanisms introduced so far, without appeal to a lexical rule or any other additional device. By incorporating adjuncts into the list of dependents of the verb, and by defining both SLASH amalgamation and the introduction of *gap-ss* elements on this level as well, we provide a uniform account for both complement and adjunct extraction. That is, to generate examples such as (72a) and (72b), we may assume that the lexical entry for *visits* allows the instantiation shown in (76), and thus may appear in a tree like (77).

(76)
$$\begin{bmatrix} \text{I-FORM} & \textit{visits} \\ \text{SUBJ} & \left\langle \boxed{1}\begin{bmatrix}\text{LOC} & \text{NP}\\ \text{SLASH} & \boxed{2}\end{bmatrix}\right\rangle \\ \text{COMPS} & \left\langle \boxed{3}\begin{bmatrix}\text{LOC} & \text{NP}\\ \text{SLASH} & \boxed{4}\end{bmatrix}\right\rangle \\ \text{DEPS} & \left\langle \boxed{1}, \boxed{3}, \begin{bmatrix}\textit{gap-ss}\\ \text{LOC} & \boxed{5}\\ \text{SLASH} & \{\boxed{5}\}\end{bmatrix}\right\rangle \\ \text{ARG-ST} & \langle \boxed{1}, \boxed{3} \rangle \\ \text{SLASH} & \boxed{2} \cup \boxed{4} \cup \{\boxed{5}\} \end{bmatrix}$$

(77)



The structure in (76) arises when, in addition to the arguments of the verb, DEPS contains an adverbial element licensed by Argument Structure Extension, and this adverbial *synsem* is further specified as a *gap-ss* by Argument Realization. The SLASH value of the *gap-ss* is incorporated into the SLASH value of the verb because SLASH amalgamation is defined in terms of the DEPS list. The example thus involves three constraints applying in conjunction with the same lexical entry.

Further evidence for relating extractibility of adjuncts to their status as complements comes from the fact that adverbials that only occur as preverbal

modifiers cannot be extracted. That is, the fronted elements in the examples in (78) cannot be interpreted as VP modifiers of the lower clause.

(78) a.*Almost, I think Kim __ found the solution.

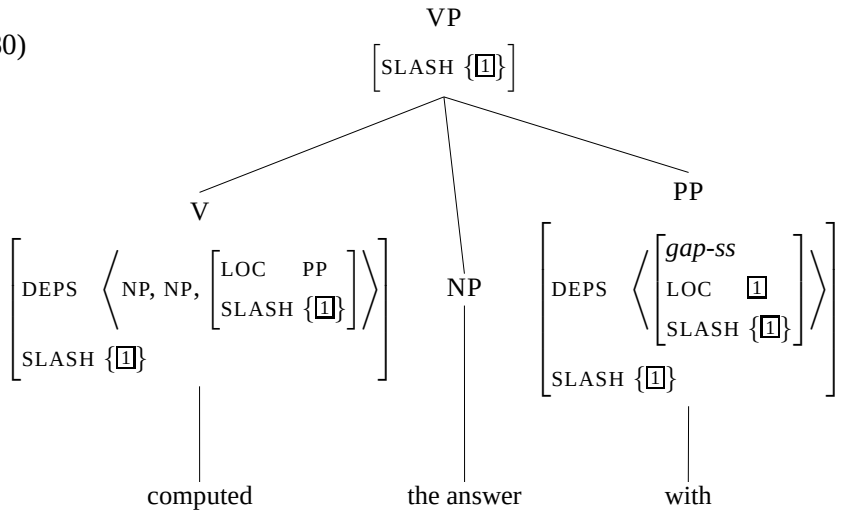
b.*Never did Kim claim that Sandy __ sang for her.

This is precisely as predicted by our unified extraction analysis: in order to be extractable, an adverbial must be a dependent of (i.e. selected by) the verb that is it modifies (semantically).

If we do not impose any further constraints, extraction out of adjuncts, as in (79), is predicted to be grammatical. This follows from the fact that the SLASH value of any adjuncts on DEPS is amalgamated by the verb. So, if the the lexical head of an adjunct phrase is slashed, this information is passed on to the verb of which it is a dependent, as shown in (80).

(79) What do you think Robin computed the answer with?

(80)



It is also straightforward to restrict such extraction to provide a treatment of adverbials as 'weak islands' (see Postal, pear). This can be accomplished simply by specifying that the relevant constituent's SLASH value must be 'set(NP)', which allows only NP extraction dependencies to be transmitted across that constituent.

Another putative piece of evidence for a configurational distinction between arguments and adjuncts is the fact that there appears to be a difference in acceptability between argument and adjunct extraction in certain contexts involving, for example, factive verbs or negation:

(81) a. What does Robin regret that Kim ate?

b.*Why does Robin regret Kim refused the offer?

c. What did you deny that John ate?

d.*How did you deny that John repaired the sink?

e. What do you doubt Sandy will like?

f.*How do you doubt Sandy will like her coffee?

g. Which soft drink don't you think Kim likes?

h.*How don't you think Kim likes her coffee?

However, part of the explanation may have to do with special properties of the adverbs *how* and *why* and phrases containing them. PS-94 present examples involving phrasal adverbials, such as those in (82) below:

(82) a. When their parents are in town next week, I doubt that the twins will attend any lectures.

b. During my term as University President, I deny there were any illegitimate appropriations of government money.

These are indisputable examples of extraction out of the putative island contexts. Hence these configurations cannot simply be declared islands for adjunct extraction.[26]

Another familiar observation that casts doubt on any solely configurational account of the examples in (81) is the fact that in these contexts adverbs which are uncontroversially subcategorized for appear to exhibit exactly the same behavior as modifying adverbs:

(83) a.*How don't you think John behaved?

b.*How did Kim doubt Sandy worded the letter?

c.*How well did Lee not think Sandy treated the animals?

On the other hand, the general strategy adopted here is compatible with a semantic account of this phenomenon, such as has been proposed by Szabolcsi and Zwarts (1993). Indeed, by making SLASH amalgamation head-driven, it seems that we provide the correct setting for a theory which imposes semantic constraints, associated with particular lexical items, on the elements within the value of SLASH.

[26] Given recent psycholinguistic research, e.g. that of Kluender (1998), it may in fact be the case that many island phenomena thought to be the consequence of grammatical constraints can be explained in terms of processing factors.

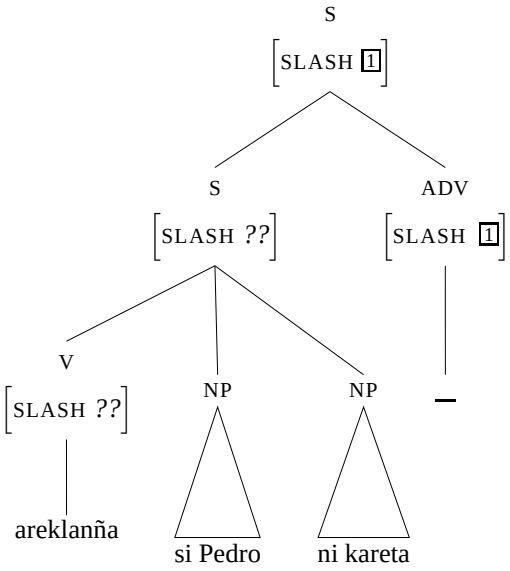
Finally, there is one more significant respect in which our analysis makes different predictions from accounts based on the traditional notion of adjuncts as modifiers. Consider the following Chamorro example:

(84) Taimänu malago'mu pära areklanña si Pedro ni kareta
how WH[OBL].want FUT WH[OBL].fix Pedro OBL car

'How do you want Pedro to fix the car?'

The embedded verb *areklanña* 'fix' bears morphology registering the fact that the manner adverb *taimänu* 'how' is extracted. Under the traditional analysis, which would assign this example the structure in (85), the verb has no access to the adverbial trace without otherwise unmotivated machinery of some kind.

(85)



However, if manner adverbs are dependents of the verb, the agreement in (83) follows directly from the account of *wh* agreement presented in section 3. Adverbial *wh* agreement parallels that of complements, involving lexical specifications like the following:

(86) $\begin{bmatrix} \text{I-FORM } \textit{areklanña} \\ \text{DEPS} \quad \left\langle \text{NP} , \text{NP} , \begin{bmatrix} \text{LOC} \quad \text{ADV} \\ \text{SLASH } \textit{neset} \end{bmatrix} \right\rangle \end{bmatrix}$

Our lexicalized approach to adjunct extraction provides a natural account, we believe, of the known facts about extraction in Chamorro and other languages which register adjunct extraction lexically and hence present a serious challenge to previous GPSG and HPSG accounts that rely on traces and the traditional distinction between complements and adjuncts.

6. Subject Extraction

6.1. ENGLISH

As noted in the introduction, earlier work in GPSG/HPSG proposed a non-uniform gapless account subject extraction. Matrix subject extraction like that in (87) was not treated as extraction at all, but rather as a simple sentential structure containing NP and VP.

(87) Who visits Alcatraz?

PS-94's account of the extraction of embedded subjects relies on the following Subject-Extraction Lexical Rule:[27]

(88) Subject Extraction Lexical Rule (SELR)

$$\begin{bmatrix} \textit{word} \\ \text{COMPS} \ \ldots \left\langle \text{S}\left[\textit{fin,unm}\right] \right\rangle \ldots \\ \text{SLASH} \ \{\ \} \end{bmatrix} \mapsto \begin{bmatrix} \text{COMPS} \ \ldots \left\langle \text{VP} \begin{bmatrix} \textit{fin} \\ \text{SUBJ}\,|\,\text{LOC}\ \boxed{3} \\ \text{SLASH}\ \{\ \} \end{bmatrix} \right\rangle \ldots \\ \text{SLASH} \ \{\boxed{3}\} \end{bmatrix}$$

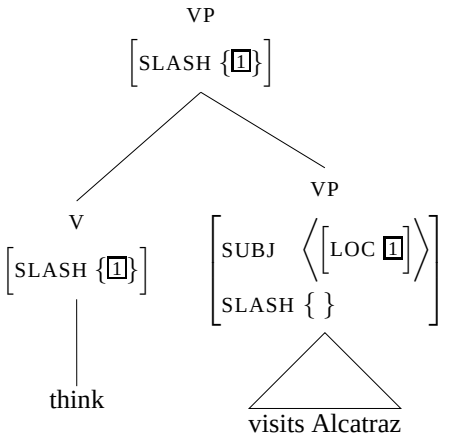
This rule produces a new lexical entry for every word whose basic entry selects for an unmarked, finite S complement. For example, (89a) would be mapped to (89b).

(89) a. $\begin{bmatrix} \text{I-FORM}\ \textit{think} \\ \text{COMPS}\ \langle\ \text{S}\ \rangle \\ \text{SLASH}\ \{\ \} \end{bmatrix}$ b. $\begin{bmatrix} \text{I-FORM}\ \textit{think} \\ \text{COMPS}\ \left\langle \text{VP}\left[\text{SUBJ}\,|\,\text{LOC}\ \boxed{1}\right] \right\rangle \\ \text{SLASH}\ \{\boxed{1}\} \end{bmatrix}$

Lexical entries like (89b) then combine directly with an unslashed VP complement in examples like (90):

[27] This is essentially a recasting of the 'Finite VP Metarule' proposed originally by Gazdar (1981).

(90)



Because the SELR applies only to those elements that COMPS-select an unmarked S complement, there is no way in this analysis to derive so called '*that*-trace violations' like (91).

(91)*Who do you think that visits Alcatraz?

The problem we noted with this analysis, apart from its lack of uniformity, is that it does not treat extracted subject arguments as gaps. Although there is no extraction registration phenomenon in English that would speak directly against this aspect of the PS-94 analysis, the cross-linguistic evidence is nonetheless highly telling. As Hukari and Levine (1996a, 1996b) have repeatedly emphasized (and as noted above), the fact that missing subjects are treated on a par with other extracted elements in all known SLASH-sensitive phenomena suggests that the subject of *visits* in (87) and (90) should be analyzed as an argument of type *gap-ss*.

The constraint-based analysis of extraction presented in section 3 allows subject arguments to be realized as *gap-ss* elements, and thus offers an interesting alternative to the account of PS-94—one which does not need to posit lexical rules or new types of grammatical constructions and which treats all verbs whose dependents are extracted uniformly.

To see this, consider the constraints we have imposed on verbs: these guarantee that verbs always have one element on their SUBJ list and that the verb's SLASH value is the union of the SLASH values of the verb's dependents. Argument Realization requires that the first element of a verb's DEPS list always be the subject, but nothing prevents this element from being of type *gap-ss*. That is, (92b) is completely consistent with the lexemic information in (92a) and all of the constraints introduced in the previous sections.

(92) a. $\begin{bmatrix} \text{ROOT} & \textit{visit} \\ \text{ARG-ST} & \langle \text{NP}, \text{NP} \rangle \end{bmatrix}$

b. $$\begin{bmatrix} \textit{word} \\ \text{I-FORM}\ \textit{visits} \\ \text{SUBJ}\ \left\langle \boxed{2} \begin{bmatrix} \textit{gap-ss} \\ \text{LOC}\ \ \boxed{1}\text{NP}[\textit{nom}]_{\textit{3sg}} \\ \text{SLASH}\ \{\boxed{1}\} \end{bmatrix} \right\rangle \\ \text{COMPS}\ \langle \boxed{3}\text{NP}[\textit{acc}] \rangle \\ \text{DEPS}\ \langle \boxed{2}, \boxed{3} \rangle \\ \text{ARG-ST}\ \langle \boxed{2}, \boxed{3} \rangle \\ \text{SLASH}\ \{\boxed{1}\} \end{bmatrix}$$

We follow Ginzburg and Sag (pear) in treating all the subject extractions in (93) in terms of the type *head-filler-phrase*.

(93) a. Who __ left?

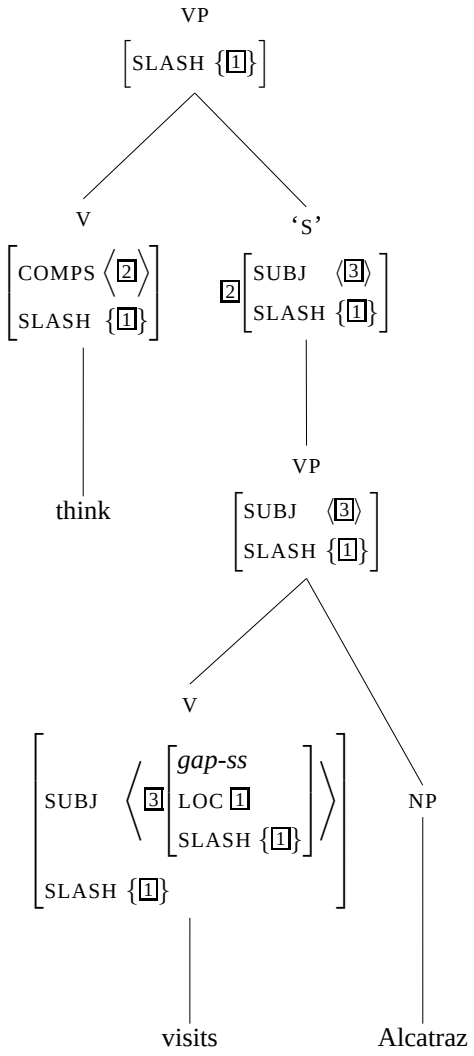
b. The person [who __ left]. . .

That is, unlike under previous PSG treatments, these clauses are not treated as instances of *head-subject-phrase*. This provides an immediate account of the fact, again first noted by Hukari and Levine (1996a, 1996b), that languages that register extraction indicate unambiguously that such clauses are instances of true extraction.

Furthermore, if a verb like *think* selects not for a saturated complement, but rather for one that is simply finite and proposition-denoting,[28] then a phrase like *visits Alcatraz* will be a possible complement of *think*. The SLASH value of *visits* will thus be passed up and amalgamated into the SLASH value of *think*, which will pass it up to the mother of the phrase *think visits Alcatraz*, as shown in (94).

[28] For detailed exposition of a theory that guarantees that proposition-denoting VPs must have a *noncanon* synsem element (e.g. a *gap-ss*) on their SUBJ list, see Ginzburg and Sag (pear).

(94)



That-trace violations like (91) can now be ruled out as a simple matter of local subcategorization. If complementizers are markers selecting a [SUBJ ⟨ ⟩] sentential complement through their SPEC features, as shown in (95), then there will be no way for a structure like (93) to cooccur with *that*.[29]

[29] Further data that can be analyzed in this way is the following contrast cited by Kayne (1984, 3):

(i) The only person who it's not essential she talk to __ is Bill.

(ii) *The only person who it's not essential __ talk to her is Bill.

Apparently, the adjective *essential*, like the complementizers *that*, *whether* and *if*, selects for a [SUBJ ⟨ ⟩] sentential complement.

(95) $\begin{bmatrix} \text{ROOT} & \textit{that} \\ \text{HEAD} & \begin{bmatrix} \text{SPEC} & \begin{bmatrix} \text{HEAD} & v \\ \text{SUBJ} & \langle\,\rangle \\ \text{COMPS} & \langle\,\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix}$

Alternatively, the same result can be achieved by treating complementizers as (CP-projecting) heads that select for a [SUBJ $\langle\,\rangle$] complement. In either case, the problem isolated by Hukari and Levine—that subjectless verbs in extraction constructions register extraction like other verbs—is neatly solved: in our analysis, all verbs whose subject is extracted bear a non-empty SLASH value, and hence are appropriately distinguished from verbs that are not part of an extraction dependency.

The analysis we have sketched here may also provide an account of the 'adverb amelioration' effect discussed by Culicover (1993) (among others). As Bresnan (1977) first observed, examples like (96) are apparently grammatical, despite being *that*-trace violations:

(96) This is the kind of person who I doubt that, under normal circumstances, would have anything to do with such a scheme.

Assuming that such examples are to be treated as grammatical,[30] they are instances of adverb 'fronting' in the embedded clause. That is, in each case the emedded clause is an instance of *head-filler-phrase* (see (23) above). One constraint on such phrases is that they must be [SUBJ $\langle\,\rangle$]. If we assume this remains true, even if the head daughter's subject is extracted,[31] then examples like (96) thus have a structure like (97).

[30] A residual worry is the decreased acceptability of examples like the following:

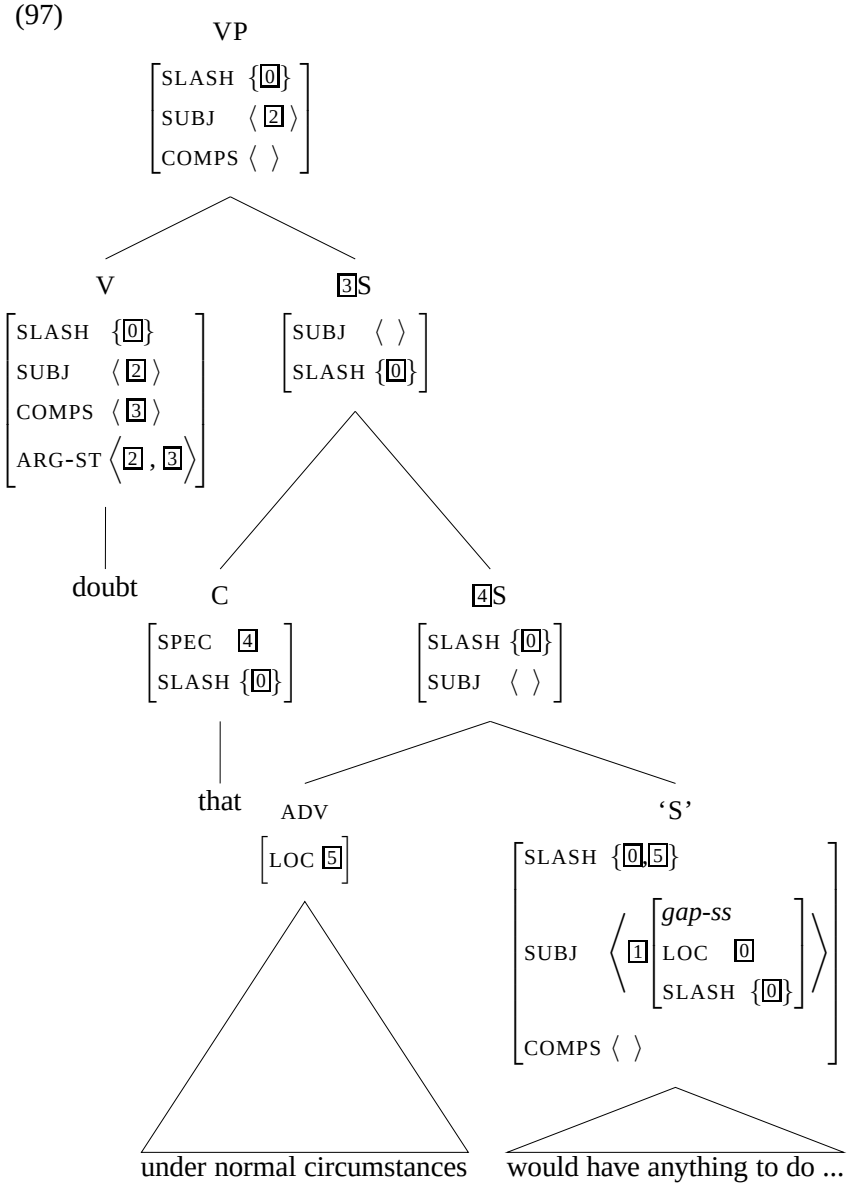
(i) ??This is the kind of person who I doubt that last year had anything to do with such a scheme.

(ii) ??Who do you think that happily visited museums?

(iii) *Which people at the conference did they think that in Paris visited museums?

[31] This may require a reformulation of the constraints on the type *head-filler-phrase* in (23) above so as to explicitly allow the head daughter's SUBJ value to be singleton or empty.

(97)



Because the embedded sentential structure is a *head-filler-phrase*, it follows that it must be [SUBJ ⟨ ⟩] and hence it is a suitable complement of the complementizer *that*.[32]

[32] Subject extraction from embedded clauses is not always possible, of course. Non-nested ('crossing') extraction dependencies like the one in (96) are allowed only under certain circumstances in English. As contrasts like the following suggest (see Fodor, 1978), the relevant factor may be whether the two fillers are of the same grammatical category (e.g. NP vs. PP).

(i) Which TV show$_i$ did you wonder how often$_j$ I saw $\_\__i$ $\_\__j$?

6.2. FRENCH

French exhibits a much-discussed contrast of complementizers that in certain respects mirrors the *that*-trace facts discussed in the previous section. The complementizer *que* is the rough equivalent of English *that*. The complementizer *qui* is not in general available as an alternative:

(98) Tu a dit *que*/**qui* cet homme est heureux.
you said that that man is happy
'You said that this man is happy.'

Yet neither *que* nor its prevocalic alternant *qu'* can combine with a clause whose subject is extracted.

(99) *L'homme que tu a dit *que*/*qu'* est heureux...
the-man that you said that is happy...
'the man that you said is happy...'

In this case, the complementizer *qui* must be used:

(100) L'homme que tu a dit *qui* est heureux...
'the man that you said that is happy...'

These contrasts receive a simple treatment in the analysis we have developed. *Que*, like English *that*, selects a phrase that is [SUBJ ⟨ ⟩]. In contrast *qui* selects a phrase that is [SUBJ ⟨*gap-ss*⟩]. It follows that the two forms are in complementary distribution. The key to the analysis is making the information about subject extraction locally accessible on the phrase selected by the complementizer.

6.3. PARASITIC GAPS

Because our analysis of SLASH values utilizes a lexical constraint defined in terms of familiar set union, it preserves the essentials of the PS-94 analysis of parasitic gaps.[33] That is, sentences like (101) are allowed because the subject and object dependents of the verb *shot* in (101a) and the object and adjunct dependent in (101b) may bear identical nonempty SLASH specifications:

(101) a. That was the rebel leader who [rivals of __] shot __.

(ii) *Which TV show$_i$ did you wonder how many students$_j$ I explained __$_i$ to __$_j$?

[33] The PS-94 analysis, which puts parasitic gaps on a par with nonparasitic gaps, runs counter to the claims made by Cinque (1990) and Postal (1994, 1998), who argue that parasitic gaps are pronominal in nature and are merely coindexed with other gaps. For a critique of this claim, see Calcagno et al. (1999).

b. Those boring old reports, Kim filed __ [without reading __].

Since SLASH amalgamation is defined in terms of set union, the verb *shot*, when selecting for a *gap-ss* object as well as an incomplete subject, may nevertheless have a SLASH value consisting of a single element:

(102)
$$\begin{bmatrix} \text{I-FORM } \textit{shot} \\ \text{SUBJ } \langle \boxed{1} \rangle \\ \text{COMPS } \langle\,\rangle \\ \text{DEPS } \left\langle \boxed{1}\begin{bmatrix} \text{LOC NP} \\ \text{SLASH } \{\boxed{2}\} \end{bmatrix}, \begin{bmatrix} \textit{gap-ss} \\ \text{LOC } \boxed{2}\text{NP} \\ \text{SLASH } \{\boxed{2}\} \end{bmatrix} \right\rangle \\ \text{SLASH } \{\boxed{2}\} \cup \{\boxed{2}\} \text{ (= } \{\boxed{2}\} \text{)} \end{bmatrix}$$

These predictions of this analysis are essentially correct, though numerous orthogonal factors, many of which we believe to be extragrammatical in nature, interact to degrade the acceptability of such sentences:

(103) a. Here is the couple that someone in our group sent [photos of __] to [relatives of __].

b. Those boring old reports, Kim filed __ [without reading __].

Furthermore, our analysis of English subject extraction solves a problem for the SELR analysis of PS-94 that was noticed by Hukari and Levine (1996a). Some constraint like the Subject Condition in (105) must be added if one insists (as PS-94 do) on blocking extractions like (104), whose acceptability status is unclear to us (Engdahl, 1983).

(104) (?)That is the only visiting dignitary that they thought my talking to __ would be inappropriate.

(105) **Subject Condition:**
A slashed subject must outrank another slashed argument.

Hukari and Levine point out, however, that subject gaps are possible in examples like (106).

(106) Robin is someone who even [$_{\text{NP}}$ friends of __] believe [$_{\text{VP}}$ should be closely watched].

On the SELR analysis, the bracketed VP in (106) is not slashed. Thus there is no second slashed element on the ARG-ST of *believe* that can license its slashed subject.

On the other hand, the acceptability of (106) follows directly from the analysis presented here. The VP complement of *believe*, like the VP complement of *think* in (94), has a nonempty SLASH value on our analysis; hence the Subject Condition will license the slashed subject of *believe*. Hukari and Levine's central point, that the SELR analysis of subject gaps was misguided, is therefore in agreement with the proposals made here.

7. Conclusion

The SLASH-based analysis of long-distance dependency constructions pioneered by Gazdar (1981) has proven to be highly resilient. The HPSG version of this analysis presented in chapter 4 of PS-94 replaces some of the mechanisms of the original theory by principles that are more readily compatible with modern feature-based approaches to grammar, but does not change the essentials of the original proposal. The ultimate version of their analysis, however, presented in the final chapter of that book and explored further in Sag and Fodor (1994) and Sag (1997), does deviate from earlier proposals by locating much of the theory of unbounded dependencies in the lexicon. While the elimination of traces, which provides part of the motivation for this innovation, may appeal to many linguists, it remained to be shown that the traceless account of extraction is indeed a solid alternative to trace-based accounts.

In particular, previous traceless theories of extraction had two main weaknesses. First, traces were only eliminated at the cost of introducing one or more lexical rules, the status of which has raised considerable controversy. Second, as Hukari and Levine (1996a) point out, previous traceless accounts fail to provide a uniform treatment of the full range of data. In the HPSG architecture of PS-94, complements and subjects are represented by two different valence features, and adjuncts and arguments are selected by widely different mechanisms. Previous lexicalist and traceless approaches thus failed to provide a simple, unifying account of extraction phenomena.

The theory of extraction we have presented here shows that this initial assessment is ultimately incorrect. Argument Realization, a single constraint on words, accounts for complement, subject, and adjunct extraction. Obviously, this constraint-based approach to extraction remedies the first problem (the controversial status of lexical rules for extraction) as well by completely eliminating lexical rules from the grammar of extraction. Finally, the important crosslinguistic observations made by Hukari and Levine about the systematic registration of extraction dependencies in adverbial and subject extraction are properly accounted for in our analysis—for the first time, we believe, in any research tradition.

References

Abeillé, A. and D. Godard: 1997, 'The Syntax of French Negative Adverbs'. In: D. Forget, P. Hirschbühler, F. Martineau, and M.-L. Rivero (eds.): *Negation and Polarity*. Amsterdam: John Benjamins.

Abeillé, A., D. Godard, and I. Sag: 1998, 'Two Kind of Composition in French Complex Predicates'. In: E. Hinrichs, T. Nakazawa, and A. Kathol (eds.): *Complex Predicates in Nonderivational Syntax*. New York: Academic Press.

Bobaljik, J.: 1998, 'Floating Quantifiers: Handle with Care'. *Glot International* **3**, 3–10.

Bohas, G.: 1990, *The Arabic Linguistic Tradition*. London: Routledge.

Borsley, R. D.: 1989a, 'An HPSG Approach to Welsh'. *Journal of Linguistics* **25**, 333–354.

Borsley, R. D.: 1989b, 'Phrase Structure Grammar and the *Barriers* Conception of Clause Structure'. *Linguistics* **27**, 843–863.

Bouma, G., R. Malouf, and I. A. Sag: 1998, 'Adjunct Scope and Complex Predicates'. Paper presented at the 20th Annual Meeting of the DGfS. Halle. To appear.

Bresnan, J.: 1977, 'Variables in the Theory of Transformations'. In: A. Akmajian, P. Culicover, and T. Wasow (eds.): *Formal Syntax*. New York: Academic Press.

Bresnan, J.: 2000, *Lexical Functional Syntax*. Oxford: Blackwell.

Calcagno, M., T. Hukari, and R. Levine: 1999, 'Parasitic Gaps in English: Non-nominal, Non-pronominal, and Case-consistent'. Ohio State University and University of Victoria.

Chae, H.-R.: 1992, 'Lexically Triggered Unbounded Discontinuities in English'. Ph.D. thesis, The Ohio State University, Columbus.

Chung, S.: 1982, 'Unbounded Dependencies in Chamorro Grammar'. *Linguistic Inquiry* **13**, 39–77.

Chung, S.: 1994, '*Wh*-agreement and "Referentiality" in Chamorro'. *Linguistic Inquiry* **25**, 1–44.

Chung, S.: 1998, *The Design of Agreement: Evidence from Chamorro*. University of Chicago Press.

Cinque, G.: 1990, *Types of $\bar{A}$-Dependencies*. Cambridge: MIT Press.

Clements, G. N.: 1984, 'Binding Domains in Kikuyu'. *Studies in the Linguistic Sciences* **14**, 37–56.

Collins, C.: 1997, *Local Economy*. MIT Press.

Copestake, A., D. Flickinger, C. Pollard, and I. A. Sag: 1999, 'Minimal Recursion Semantics: An Introduction'. Stanford University and Ohio State University.

Culicover, P.: 1993, 'Evidence against ECP Accounts of the That-*t* Effect'. *Linguistic Inquiry* **24**, 557–561.

Davis, T.: 1996, 'Lexical Semantics and Linking in the Hierarchical Lexicon'. Ph.D. thesis, Stanford University.

Diesing, M.: 1990, 'Verb Movement and the Subject Position in Yiddish'. *Natural Language and Linguistic Theory* **8**, 41–79.

Dowty, D. R. and B. Brodie: 1984, 'The Semantics of "Floated" Quantifiers in a Transformationless Grammar'. In: R. Hendrick (ed.): *Proceedings of the West Coast Conference on Formal Linguistics 3*. Stanford.

Dukes, M.: 1992, 'On the Status of Chamorro *Wh*-agreement'. In: J. Mead (ed.): *Proceedings of the West Coast Conference on Formal Linguistics 11*. Stanford.

Engdahl, E.: 1983, 'Subject Gaps: an Asymmetry between Questions and Relatives in Norweigan'. In: C. Jones and P. Sells (eds.): *Papers from the Fourteenth Annual Meeting of the North Eastern Linguistic Society*.

Flickinger, D. and J. Nerbonne: 1992, 'Inheritance and Complementation: A Case Study of *Easy* Adjectives and Related Nouns'. *Computational Linguistics* **18**, 269–309.

Fodor, J.: 1978, 'Parsing Strategies and Constraints on Transformations'. *Linguistic Inquiry* **9**, 427–473.

Foley, W. A. and R. D. Van Valin: 1984, *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.

Gazdar, G.: 1981, 'Unbounded Dependencies and Coordinate Structure'. *Linguistic Inquiry* **12**, 155–184.

Gazdar, G., E. Klein, G. Pullum, and I. A. Sag: 1985, *Generalized Phrase Structure Grammar*. Cambridge: Harvard University Press.

Gazdar, G., E. Klein, G. K. Pullum, and I. A. Sag: 1984, 'Foot Features and Parasitic Gaps'. In: W. Geest and Y. Putseys (eds.): *Sentential Complementation.* Dordrecht: Foris.

Georgopolous, C.: 1985, 'Variables in Palauan Syntax'. *Natural Language and Linguistic Theory* **3**, 59–94.

Gibson, E. and G. Hickok: 1993, 'Sentence Processing With Empty Categories'. *Language and Cognitive Processes* **8**, 147–161.

Ginzburg, J. and I. A. Sag: to appear, *English Interrogative Constructions*. Stanford: CSLI Publications.

Goodall, G.: 1987, *Parallel Structures in Syntax: Coordination, Causatives, and Restructuring*. Cambridge and New York: Cambridge University Press.

Gorrell, P.: 1993, 'Evaluating the Direct Association Hypothesis: A Reply to Pickering and Barry'. *Language and Cognitive Processes* **8**, 129–146.

Haik, I.: 1990, 'Anaphoric, Pronominal and Referential INFL'. *Natural Language and Linguistic Theory* **8**, 347–374.

Hinrichs, E. and T. Nakazawa: 1994, 'Linearizing AUXs in German Verbal Complexes'. In: J. Nerbonne, K. Netter, and C. Pollard (eds.): *German in Head-driven Phrase Structure Grammar*. Stanford: CSLI Publications, pp. 11–38.

Hinrichs, E. and T. Nakazawa: 1996, 'Applying Lexical Rules Under Subsumption'. In: *Proceedings of the 16th International Conference on Computational Linguistics (COLING)*. Copenhagen, pp. 543–549.

Höhle, T.: 1995, 'The Complement Extraction Lexical Rule and Variable Argument Raising'. Paper presented at HPSG Workshop. Universität Tübingen.

Hukari, T. E. and R. D. Levine: 1995, 'Adjunct Extraction'. *Journal of Linguistics* **31**(2), 195–226.

Hukari, T. E. and R. D. Levine: 1996a, 'Phrase Structure Grammar: the Next Generation'. *Journal of Linguistics* **32**, 465–496.

Hukari, T. E. and R. D. Levine: 1996b, 'Subject Extraction'. Paper presented at Third International Conference on Head-driven Phrase Structure Grammar, Marseille.

Kaplan, R. M. and A. Zaenen: 1989, 'Long-distance Dependencies, Constituent Structure and Functional Uncertainty'. In: M. R. Baltin and A. S. Kroch (eds.): *Alternative Conceptions of Phrase Structure*. University of Chicago Press.

Kathol, A.: 1995, 'Linearization-Based German Syntax'. Ph.D. thesis, The Ohio State University.

Kayne, R.: 1980, 'Extensions of binding and case-marking'. *Linguistic Inquiry* **11**, 55–91. Reprinted in Kayne (1984).

Kayne, R.: 1984, *Connectedness and Binary Branching*. Dordrecht: Foris.

Kayne, R. and J.-Y. Pollock: 1978, 'Stylistic Inversion, Successive Cyclicity, and Move NP in French'. *Linguistic Inquiry* **9**, 93–133.

Kim, J.-B. and I. A. Sag: 1995, 'The Parametric Variation of English and French Negation'. In: J. Camacho, L. Choueiri, and M. Watanabe (eds.): *Proceedings of the West Coast Conference on Formal Linguistics 14*.

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*, Vol. 29 of *Syntax and Semantics*. San Diego: Academic Press, pp. 241–279.

Koenig, J.-P.: 1994, 'Lexical Underspecification and the Syntax/Semantics Interface'. Ph.D. thesis, University of California, Berkeley.

Koenig, J.-P.: 1999, *Lexical Relations*. Stanford: CSLI Publications.

Kroeber, P.: 1997, 'Relativization in Thompson River Salish'. *Anthropological Linguistics* **39**, 376–422.

Kroeger, P.: 1993, *Phrase Structure and Grammatical Relations in Tagalog*. Stanford: CSLI Publications.

Maling, J.: 1989, 'Adverbials and Structural Case in Korean'. In: S. Kuno, I.-H. Lee, J. Whitman, S.-Y. Bak, Y.-S. Kang, and Y.-J. Kim (eds.): *Harvard Studies in Korean Linguistics III*. Cambridge, MA: Harvard University, pp. 297–308.

Maling, J.: 1993, 'Of Nominative and Accusative: the Hierarchical Assignment of Grammatical Case in Finnish'. In: A. Holmberg and U. Nikanne (eds.): *Case and other functional categories in Finnish syntax*. Dordrecht: Mouton de Gruyter, pp. 51–76.

Maling, J. and A. Zaenen: 1978, 'The Non-universality of a Surface Filter'. *Linguistic Inquiry* **9**, 475–497.

Malouf, R.: 1999, 'West Greenlandic Noun Incorporation in a Monohierarchical Theory of Grammar'. In: G. Webelhuth, A. Kathol, and J.-P. Koenig (eds.): *Lexical and Constructional Aspects of Linguistic Explanation*. Stanford: CSLI Publications, pp. 47–62.

Manning, C.: 1996, *Ergativity: Argument Structure and Grammatical Relations*. Stanford: CSLI Publications.

Manning, C. and I. A. Sag: 1998, 'Argument Structure, Valence, and Binding'. *Nordic Journal of Linguistics* **21**, 107–144.

Manning, C., I. A. Sag, and M. Iida: 1999, 'The Lexical Integrity of Japanese Causatives'. In: R. Levine and G. Green (eds.): *Readings in HPSG*. Cambridge: Cambridge University Press, pp. 39–79.

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

McCloskey, J.: 1989, 'Resumptive Pronouns, $\bar{\text{A}}$-binding, and Levels of Representations in Irish'. In: R. Hendrick (ed.): *Syntax and Semantics 23: The syntax of the modern Celtic languages*. New York: Academic Press, pp. 199–248.

McConnell-Ginet, S.: 1982, 'Adverbs and Logicalc Form'. *Language* **58**, 144–184.

Melčuk, I.: 1979, *Studies in Dependency Syntax*. Ann Arbor: Karoma Publishers.

Miller, P.: 1992, *Clitics and Constituents in Phrase Structure Grammar*. New York: Garland.

Miller, P. and I. A. Sag: 1997, 'French Clitic Movement Without Clitics or Movement'. *Natural Language and Linguistic Theory* **15**, 573–639.

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

Müller, S.: 1996, 'Complement Extraction Lexical Rules and Argument Attraction'. In: D. Gibbon (ed.): *Natural Language Processing and Speech Technology. Results of the 3rd KONVENS Conference, Bielefeld, October 1996*. Berlin, New York, pp. 223–236.

Munn, A.: 1992, 'A Null Operator Analysis of Parasitic Gaps'. *The Linguistic Review* **19**, 1–26.

Pickering, M. and G. Barry: 1991, 'Sentence Processing without Empty Categories'. *Language and Cognitive Processes* **6**(3).

Pollard, C. and I. A. Sag: 1987, *Information-based Syntax and Semantics*. Stanford: CSLI Publications.

Pollard, C. and I. A. Sag: 1994, *Head-driven Phrase Structure Grammar*. Stanford: CSLI Publications.

Postal, P. M.: 1994, 'Parasitic and Pseudoparasitic Gaps'. *Linguistic Inquiry* **25**, 63–117.

Postal, P. M.: 1998, *Three Investigations of Extraction*. Cambridge, MA: MIT Press.

Postal, P. M.: to appear, 'Islands'. In: M. Baltin and C. Collins (eds.): *The Handbook of Syntactic Theory*. Oxford: Blackwell.

Przepiórkowski, A.: 1999a, 'Case assigment and the complement/adjunct dichotomy: A non-configurational constraint-based approach'. Ph.D. thesis, Universität Tübingen.

Przepiórkowski, A.: 1999b, 'On Case Assignment and 'Adjuncts as Complements". In: G. Webelhuth, A. Kathol, and J.-P. Koenig (eds.): *Lexical and Constructional Aspects of Linguistic Explanation*. Stanford: CSLI Publications, pp. 231–246.

Przepiórkowski, A.: 1999c, 'On Complements and Adjuncts in Polish'. In: R. Borsley and A. Przepiórkowski (eds.): *Slavic in Head-Driven Phrase Structure Grammar*. Stanford: CSLI Publications.

Reape, M.: 1994, 'Domain Union and Word Order Variation in German'. In: J. Nerbonne, K. Netter, and C. Pollard (eds.): *German in Head-driven Phrase Structure Grammar*. Stanford: CSLI Publications, pp. 151–197.

Ross, J. R.: 1967, 'Constraints on Variables in Syntax'. Ph.D. thesis, MIT.

Rothstein, S.: 1991, 'Heads, Projections, and Categorial Determination'. In: K. Leffel and D. Bouchard (eds.): *Views on Phrase Structure*. Dordrecht: Kluwer.

Sag, I. A.: 1997, 'English Relative Clause Constructions'. *Journal of Linguistics* **33**, 431–484.

Sag, I. A.: 1998, 'Without a Trace'. Unpublished manuscript, Stanford University.

Sag, I. A. and J. D. Fodor: 1994, 'Extraction Without Traces'. In: *Proceedings of the West Coast Conference on Formal Linguistics 13*. Stanford.

Sells, P.: 1984, 'Syntax and Semantics of Resumptive Pronouns'. Ph.D. thesis, University of Massachusetts, Amherst.

Sportiche, D.: 1988, 'A Theory of Floating Quantifiers and Its Corrolaries for Constituent Structure'. *Linguistic Inquiry* **19**, 425–449.

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

Szabolcsi, A. and F. Zwarts: 1993, 'Weak Islands and an Algebraic Semantics for Scope Taking'. *Natural Language Semantics* **2**, 1–50.

Tesnière, L.: 1959, *Elements de Syntaxe Structurale*. Paris: C. Klincksieck.

Torrego, E.: 1984, 'On Inversion in Spanish and Some of its Effects'. *Linguistic Inquiry* **15**, 103–129.

van Noord, G. and G. Bouma: 1994, 'Adjuncts and the Processing of Lexical Rules'. In: *Proceedings of the 15th International Conference on Computational Linguistics (COLING)*. Kyoto, pp. 250–256.

van Noord, G. and G. Bouma: 1996, 'Dutch Verb Clustering without Verb Clusters'. In: P. Blackburn and M. de Rijke (eds.): *Specifying Syntactic Structures*. Stanford: CSLI Publications, pp. 123–153.

Webelhuth, G.: 1992, *Principles and Parameters of Syntactic Saturation*. New York: Oxford University Press.

Wechsler, S.: 1995, *The Semantic Basis of Argument Structure*. Stanford: CSLI Publications.

Wechsler, S. and Y.-S. Lee: 1996, 'The Domain of Direct Case Assignment'. *Natural Language and Linguistic Theory* **14**, 629–664.