

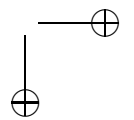
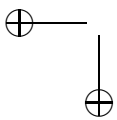
Features:

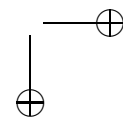
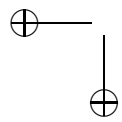
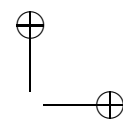
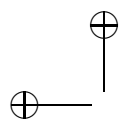
Perspectives on a Key Notion in Linguistics

Anna Kibort and Greville G. Corbett

CLARENDON PRESS • OXFORD

2009





Contents

List of figures vi

List of tables vii

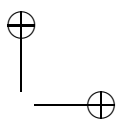
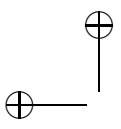
1 Feature Geometry and Predictions of Locality 1

Ivan A. Sag

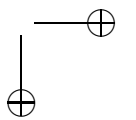
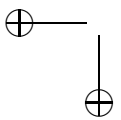
- 1.1 Introduction 1
- 1.2 Background 2
- 1.3 Locality of Construction 15
- 1.4 Sign-Based Construction Grammar 17
- 1.5 Predictions of Locality 22
- 1.6 Some Analytic Challenges 27
- 1.7 Conclusion 38

List of figures

1.1	An Illicit Selection of an Embedded NP	4
1.2	Kajita’s (1968) Clause Structure	6
1.3	Subcategorization by Subjunctive Clause (Kajita)	7
1.4	HEAD Features in GPSG	8
1.5	Lexical ID Rules in GPSG	9
1.6	Head Feature Identity within a Headed Complement (GPSG)	10
1.7	Head and Foot Feature Identity within a VP Complement (GPSG)	11
1.8	Feature Geometry of Pollard and Sag 1994	16
1.9	A Tag-Question	17
1.10	The Sign in SBCG	18
1.11	A Clausal Construct	19
1.12	Type Declarations	19
1.13	A SBCG Type Hierarchy	20
1.14	Interaction of the Subcategorization and Head Feature Principles	24
1.15	A <i>For-To</i> Clause	28
1.16	A Filler-Gap Dependency in SBCG	30
1.17	Case Government in Martuthunira	32
1.18	XARG Analysis of Genitive-Embedding NP	35
1.19	XARG Analysis of <i>your fancy</i>	35
1.20	Literal and Idiomatic Strings Relations	38



List of tables



1

Feature Geometry and Predictions of Locality

IVAN A. SAG

1.1 Introduction

This chapter deals with a number of issues having to do with locality in natural language.¹ **Locality of selection** is the problem of delimiting what syntactic and semantic information lexical items select. Related issues include the proper analysis of idiomatic expressions, control of overt pronominals, and cross-linguistic variation in lexical sensitivity to filler-gap dependencies. Closely related to selectional locality is the issue of **locality of construction** – the problem of delimiting the syntactic and semantic information accessible to grammar rules. These issues have considerable history in the field, though matters of locality are sometimes left implicit in theoretical discussions.

After providing some necessary background, I will propose a version of grammatical theory that embodies a particular hypothesis about locality. In the general theory I outline, the feature geometry serves to delimit the grammatical information accessible for lexical selection or constructional constraints.

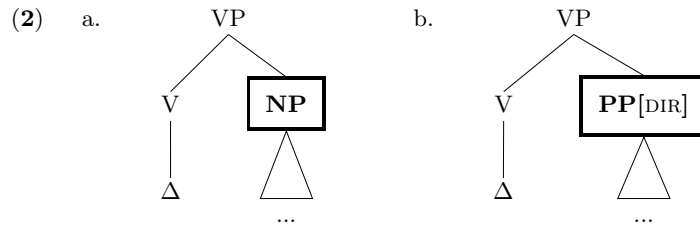
¹ For valuable discussions about locality, I would like to thank Emily Bender, Ann Copestake, Grev Corbett, Bill Croft, Bruno Estigarribia, Charles Fillmore, Dan Flickinger, Adele Goldberg, Andreas Kathol, Paul Kay, Bob Levine, Detmar Meurers, Laura Michaelis, Carl Pollard, Jan Strunk, and Tom Wasow. I am particularly grateful to Detmar Meurers and Stefan Müller for detailed comments on an earlier draft of this chapter and to Grev Corbett and Anna Kibort for excellent editorial advice. This work was supported in part by grant BCS-0094638 from the National Science Foundation to Stanford University and in part by the Research Collaboration between NTT Communication Science Laboratories, Nippon Telegraph and Telephone Corporation and CSLI, Stanford University.

1.2 Background

The locality of selection is one of the theoretical issues that were hotly debated during the 1960s. For example, Chomsky (1965, Ch. 2) proposed that the lexical entries of verbs and other lexical ‘formatives’ include ‘strict subcategorization restrictions’ like those shown in (1):²

- (1) a. prove, V, [+ __ NP]
 b. run, V, [+ __ DIR]

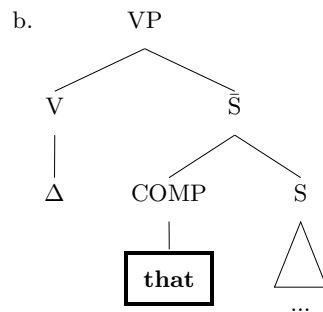
Context-sensitive lexical insertion transformations (which involved the substitution of a lexical formative for a dummy symbol ‘ Δ ’) were subject to a ‘matching condition’ that required the subcategorization restrictions to match the local context in the deep structure phrase marker. Chomsky proposed that the matching condition obeyed a principle of ‘strict locality’, which stipulated that strict subcategorization restrictions like those illustrated in (1) could only make reference to (could only be matched against) elements that are dominated by the VP directly dominating the V in deep structure subtrees like (2a,b):



Strict locality imposed an upper bound on the domain of subcategorization, but not a lower bound. That is, an element referred to by a subcategorization restriction did not have to be a sister of the V; it could be an element embedded within a sister of the V, as in (3):

- (3) a. believe, V, [+ __ that S]

² The field has fallen into an oddly mutated use of the verb *subcategorize*. One frequently finds in the literature expressions like ‘This verb subcategorizes for X’ or ‘This verb is subcategorized for X’. Here and throughout, I will maintain what I believe is the original way of describing the dependencies in question, e.g. ‘This verb is subcategorized by X. That is, the particular syntactic environment X is used to classify the verb in question into the given subcategory.’



But strict locality sharply distinguished subcategorization restrictions from selectional restrictions, the similar device introduced by Chomsky to analyze semantic cooccurrence restrictions. The selectional restrictions of a verb, for example, were permitted to access properties of the subject NP, but the strict subcategorization restrictions were not.

This matter was taken up anew by Kajita (1968), who argued that Chomsky’s notion of strict locality was both too strong and too weak. In particular, Kajita (p. 96) argued, on the basis of contrasts like (4a,b), that subcategorizational domains should be extended to include a verb’s subject:³

- (4) a. That Kim was right bothered me.
 b. *That Kim was right loved me.

Although contrasts like this might be explained away as semantic (selectional) in nature, there are other minimal pairs that perhaps make Kajita’s point more convincingly:

- (5) a. The question of whether Kim was right perplexed me.
 b. *?Whether Kim was right perplexed me.

In any case, it is now well established that many languages have verbs that select a subject with idiosyncratic case properties (e.g. Icelandic verbs requiring a ‘quirky’ dative, accusative or genitive subject; see Thráinsson 1979). Hence, given that case information is (at least partly)

³ This argument, of course, turns on the assumption that English has hierarchical clause structure, and not the flat structure assumed, for example, in a number of proposals for German, Japanese and other language with considerable word order freedom. Assuming the flat structure for clauses, the subject is accessible to a verb without modifying Chomsky’s theory of strict subcategorization.

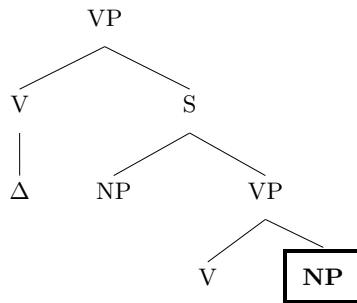


FIG. 1.1 An Illicit Selection of an Embedded NP

syntactic in nature, permitting the syntactic selection of subjects, as Kajita suggested, provides the most straightforward account of quirky subject case and related phenomena.

Chomsky’s strict locality proposal was too weak, Kajita argued, because it allowed subcategorization restrictions to access elements deeply embedded within a verb’s complement. For example, under Chomsky’s definition of strict locality, an object within a clause would be locally accessible to a verb that selected that clause as a complement, as in Figure 1.1. The objection runs as follows: although we commonly find verbs like *prove*, which require a direct object NP (and disallow a PP complement), there are no languages (as far as we know) where we find a verb like *prove* that imposed the same requirement on the complementation pattern realized within its sentential complement. That is we would not expect to find a verb *evorped* whose selectional properties produced contrasts like the following:

- (6) a. Lee **evorped** that someone bought the car.
- b.*Lee **evorped** that someone died —.
- c.*Lee **evorped** that someone ran into the room.

Kajita is to my knowledge the first to point out the theoretical importance of characterizing the lower bound on subcategorization restrictions.

Kajita also considered examples like the following, arguing that the verb *serve* requires an infinitival VP complement (an S, in his system) that contains a direct object NP:

- (7) a. The ice served to chill the beer.
 a.*The ice served to melt

To accommodate this contrast and other data he considers, Kajita (1968: 105) proposed that the upper bound of a verb’s subcategorization domain be the minimal S node that dominates it and that the lower bound be determined by a constraint requiring that the path from the upper bound to the selected constituent contain at most one S node.⁴ Kajita’s theory must be understood in terms of the particular theory of deep structure that he was assuming, which countenanced deep phrase markers like the one shown in Figure 1.2 as the analytic basis for sentences like *John thinks that Mary is certainly smart*.

The intent here was to rule out the possibility of a verb substituted for Δ whose lexical entry contained a subcategorization restriction that made reference to, say, the Present Tense of the embedded clause or the AP within that clause’s VP.

However, Kajita treated subjunctive selection, for example in *He demanded that everyone take the examination next Monday*, in terms of selection for future tense, as illustrated in Figure 1.3 (Kajita’s deep structure for sentences like *He suggested that everyone take the exam next Monday*). Thus, under his assumptions about clausal structure (which were justified in considerable detail), it was crucial that subcategorizational domains be allowed to cross exactly one sentential node.

However, Kajita’s conclusions about the verb *serve* were reassessed by Higgins (1979: 173, nt. 5), who argued that the correct generalization is a semantic requirement: the unexpressed subject of the VP complement of the verb *serve* must be interpretable as an instrument. Shieber supported this conclusion by observing (p.c. reported in Pollard and Sag (1987, p. 145)) that examples like the following follow Higgins’ constraint, but not Kajita’s:

- (8) *Kim served to break the window with a hammer.

The verb *serve* thus imposes semantic constraints on the unexpressed subject of its VP complement, but makes no direct reference to the internal syntactic properties (e.g. the presence of an object NP) within that VP.

I am not aware of transformational studies that have sought to refine or update Kajita’s conclusions. Indeed, the question of locality of

⁴ I am loosely paraphrasing Kajita’s theory, replacing his distinction ‘width’ and ‘depth’ of the subcategorization by ‘upper’ and ‘lower bound’.

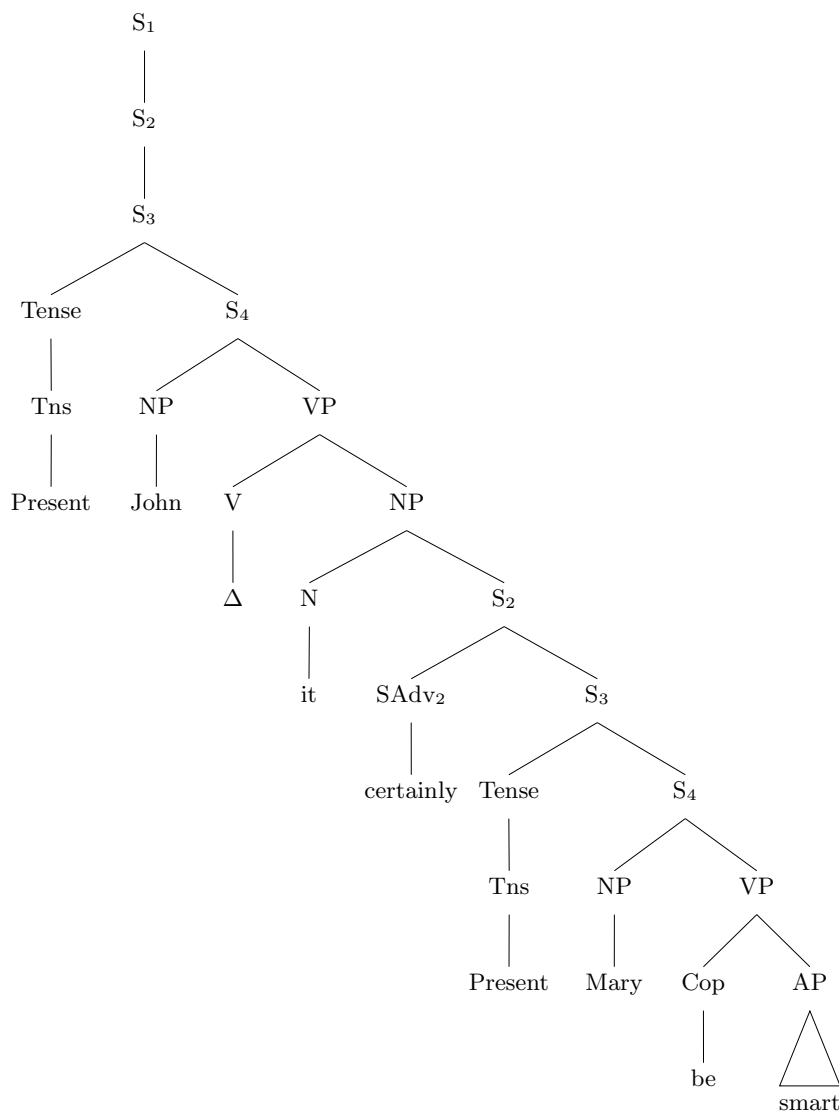


FIG. 1.2 Kajita's (1968) Clause Structure

subcategorization seems to have fallen by the wayside within the mainstream of transformational grammar. It is important to realize, however, that ‘ \bar{X} Theory’, as developed by Chomsky (1970) (but cf. Harris 1946), bears on this question. A verb that is subcategorized by an

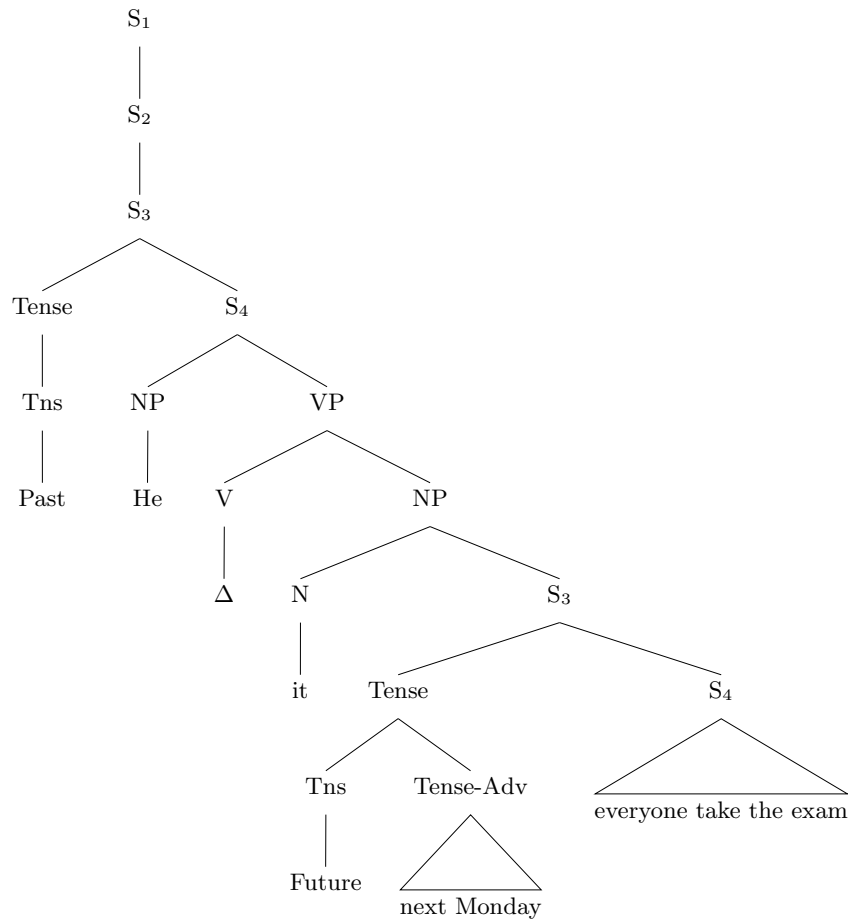


FIG. 1.3 Subcategorization by Subjunctive Clause (Kajita)

NP complement (that is, a transitive verb) really selects a phrase with a nominal head. And \bar{X} Theory, which relies on the reformulation of syntactic categories as feature structures, provides a way of projecting the category information of the lexical head ‘up’ to its maximal projection (e.g. the maximal NP headed by a given noun, the maximal AP headed by a given adjective, etc.). \bar{X} Theory thus plays a crucial role in considerations of locality – a verb can refer to the category features of the phrases it combines with, i.e. the phrases (NP, AP, etc.) that are sisters of the verb, and \bar{X} Theory will ensure that those phrases will be headed by a word of the appropriate category.

CASE (values in $\{nom, acc\}$, specified for all NPs, but distinguishing among pronouns, e.g. *she* vs. *her*),
 VFORM (values in $\{fin, base, inf, prp, psp, ger\}$ distinguishing the various inflected forms of the V that heads a VP or S),
 NFORM (values in $\{norm, it, there\}$, distinguishing referential nominals from dummies),
 PFORM (values in $\{to, of, loc, dir\}$, distinguishing the various kinds of prepositions (and PPs) that can be involved in subcategorization),
 PRED (values in $\{+, -\}$, distinguishing the predicative Xs (and XPs) from their nonpredicative counterparts),
 AUX (values in $\{+, -\}$, distinguishing the auxiliary verbs (and VPs) from their nonauxiliary counterparts), and
 SLASH (values in sets or lists of categories, distinguishing ‘saturated’ phrases from those that contain one or more unbound gaps of a particular kind).

FIG. 1.4 HEAD Features in GPSG

These ramifications of \bar{X} Theory played an important analytic role in Generalized Phrase Structure Grammar (GPSG). Gazdar (1981, 1982) and Gazdar et al. (1985) argued that \bar{X} Theory, with a slightly enriched inventory of syntactic features, provides the basis for a wholesale revision of linguistic theory, one that eliminates transformational operations altogether. GPSG researchers proposed that the ‘HEAD’ features, those whose specifications were passed up from head daughter to mother in a headed structure, included not only N and V, which (following Chomsky) were used to (coarsely) distinguish grammatical categories, but also all the features illustrated in Figure 1.4.

With this feature inventory, the explanatory domain of \bar{X} Theory is expanded to include not only the locality of category selection, but also the locality of case assignment, verb form government, selection of expletives, preposition selection, auxiliary selection, and the selection of phrases containing gaps of a particular kind (e.g. by *tough*-adjectives in English). Assuming that specifications for these features are ‘percolated up’ from lexical heads to the phrases they project (by the Head Feature Principle (HFP), an uncontroversial principle of \bar{X} Theory), the information required for the analysis of all these phenomena becomes locally accessible to the lexical elements that select those phrasal projections as complements.

In other words, once \bar{X} Theory and an expanded inventory of HEAD features were adopted, proponents of GPSG were able to reformulate

$VP \rightarrow H[1]$, where $V[1]$ is in $\{walk, die, laugh, \dots\}$
 $VP \rightarrow H[2]$, $NP[acc]$, where $V[2]$ is in $\{prove, clarify, reveal, \dots\}$
 $VP \rightarrow H[3]$, $VP[base]$, where $V[3]$ is in $\{can, should, may, \dots\}$
 $AP \rightarrow A[27]$, $VP[inf]/NP[acc]$, where $A[27]$ is in $\{tough, easy, \dots\}$
 \dots

FIG. 1.5 Lexical ID Rules in GPSG

grammar rules as shown in Figure 1.5, where verbs are subcategorized only by properties of their sister constituents:⁵

This ‘context-free’ theory of subcategorization relies on the HFP and other general principles (e.g. the Foot Feature Principle) to define the domain in which subcategorization restrictions hold, e.g. in structures like those in Figures 1.6 and 1.7. In fact, given the possibility of modification and the unbounded expansion of ‘slashed’ constituents, the domain over which subcategorization is allowed in a GPSG/HPSG approach is in principle unbounded, as it should be, given across-the-board effects in coordination, and unbounded effects in modification, extraposition, and other structures, as illustrated for VFORM selection in (9):

- (9) a. Kim will [**leave**/***leaving**/***left** home].
 b. Kim will [[**leave** home] and [**get** famous]].
 c. Kim will [apparently [never [**leave** home]]].
 d. Kim will [[[**drink** [so much]] [at the party]] [that we’ll be embarrassed]].

To put it somewhat differently, GPSG did not deny that there were long-distance dependency phenomena of the sort just illustrated. Rather, the claim made by GPSG (and also by the HPSG approach to be

⁵ Note that here the following abbreviations are used:

VF = VFORM

$$\begin{array}{l}
 V[i] = [\text{SUBCAT } i] \\
 VP[base] = \begin{bmatrix} N & - \\ V & + \\ \text{VFORM} & base \\ \text{SUBJ} & - \end{bmatrix}
 \end{array}
 \qquad
 \begin{array}{l}
 NP[acc] = \begin{bmatrix} N & + \\ V & - \\ \text{CASE} & acc \\ \text{BAR} & 2 \end{bmatrix} \\
 VP[inf]/NP = \begin{bmatrix} N & - \\ V & + \\ \text{VFORM} & inf \\ \text{SUBJ} & - \\ \text{SLASH} & NP[acc] \end{bmatrix}
 \end{array}$$

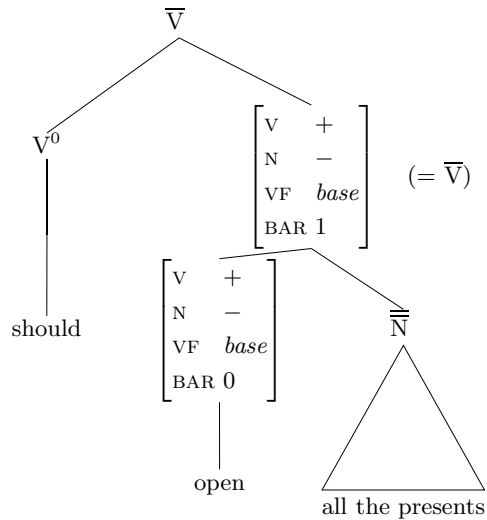
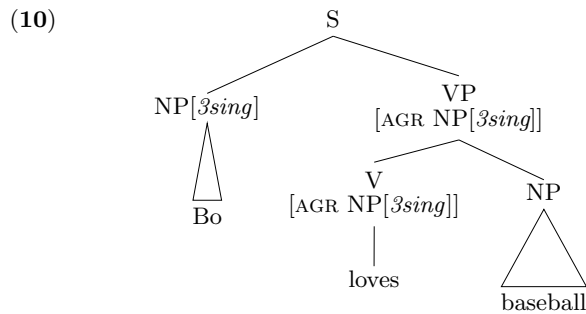


FIG. 1.6 Head Feature Identity within a Headed Complement (GPSG)

discussed below) is that non-local dependency phenomena are a consequence of strictly local constraints (e.g. lexical specifications involving the category, meaning, case, etc. of a word’s selected dependents) and their interaction with independent principles of grammar, such as the HFP.

GPSG accommodated subcategorization by subjects in terms of another HEAD feature AGR, which allowed a verb to ‘pass up’ information (again, via the HFP) to its VP projection, whose AGR value had to be identified with the subject NP, by a separate principle (the Control Agreement Principle):



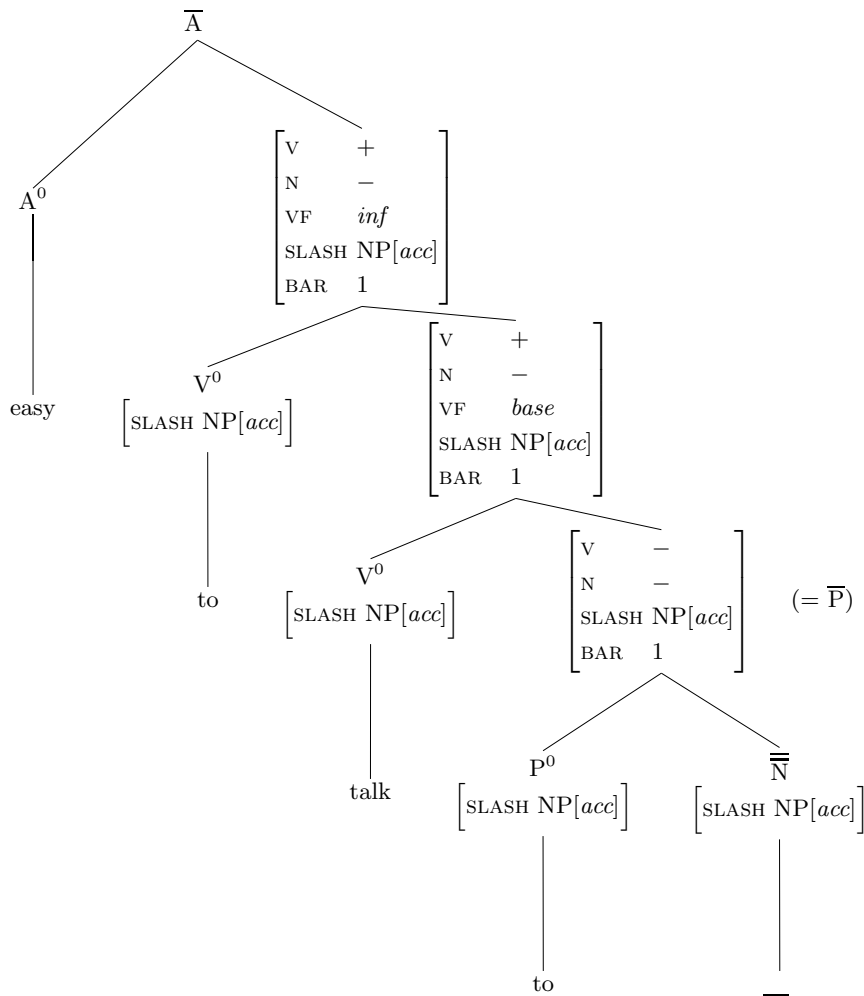
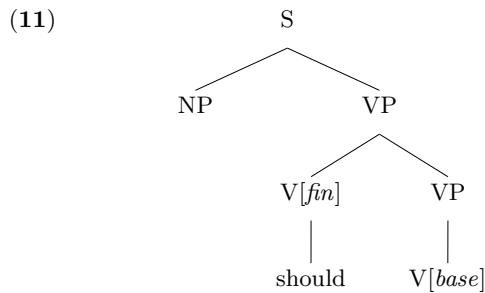


FIG. 1.7 Head and Foot Feature Identity within a VP Complement (GPSG)

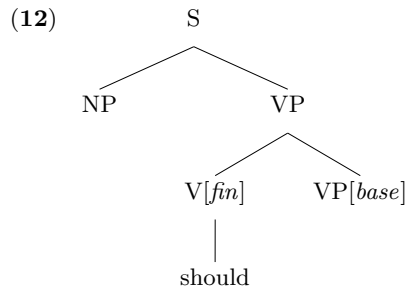
Note also that the difference between subjunctive verbs ([VFORM *base*]) and indicative verbs ([VFORM *fin*]) is projected by the HFP, and thus provides an account of Kajita’s example in Figure 1.3 above in terms of context-free subcategorization: *believe* is subcategorized by an S[*fin*] complement; *demand* is subcategorized by an S[*base*]. Without a doubt, GPSG achieved a theory of subcategorization that embodied a notion of locality quite similar to the one proposed by Kajita. The GPSG theory is

not about deep structure phrase markers, of course; GPSG embraced the ambitious goal of generating surface syntactic structures directly. And in the GPSG theory, Kajita’s domain stipulation, as well as the exceptions to it that must be countenanced in a surface-based subcategorization theory, actually follows as a theorem from the nature of the subcategorization mechanism and its interaction with independently motivated grammatical principles – a welcome result.

GPSG’s approach to subcategorization was based on local trees and the decomposition of categories via syntactic features. The best known tree-based approach to subcategorization, however, is probably Tree-Adjoining Grammar (TAG; first proposed by Joshi et al. (1975)), which differs from GPSG in grounding sentence generation not in local trees, but rather in elementary trees that can be viewed as approximating Kajita’s local domains:



In TAGs, lexically anchored elementary trees like this can undergo two kinds of operations: a tree structure can be substituted for either of the unexpanded nodes in (11) (the NP or the V[base]) or else an auxiliary tree can be grafted into the middle of (11) by the adjunction operation. In this set-up, the question of locality is in essence the question of how deep elementary trees can be. One might attempt to retain \bar{X} Theory within TAG, for example, and replace (11) with a more shallow tree like (12), possibly providing a tighter theory of locality:

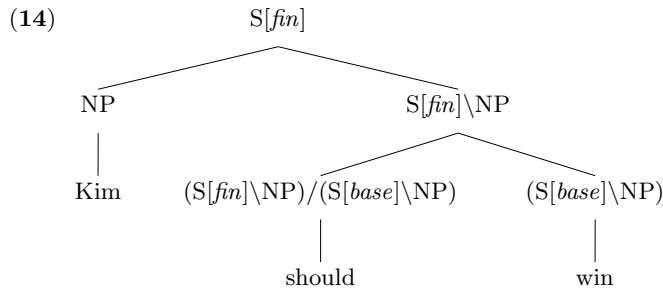


A lexically anchored tree in TAG corresponds to a lexical entry in other frameworks. In Categorical Grammar (CG) and Head-Driven Phrase Structure Grammar (HPSG), for instance, a notion of locality is built into the structure of lexical categories. The GPSG grammar rules in Figure 1.5 above correspond to lexical entries like the following, where NP, S, and AP are abbreviations for feature structure categories similar to those illustrated above:⁶

- (13)
- a. $S \backslash NP$: { *walk, die, laugh, \dots* }
 - b. $(S \backslash NP) / NP[acc]$: { *prove, clarify, reveal, \dots* }
 - c. $(S \backslash NP) / (S[base] \backslash NP)$: { *can, should, may, must, \dots* }
 - d. $(AP) / ((S \backslash NP) / NP[acc])$: { *tough, easy, \dots* }
- ...

Hence, assuming a simple regime of function application for the construction of basic sentences, as illustrated in (14), these lexical representations provide an extended locality domain for subcategorization that is, again, quite like Kajita’s:

⁶ A note on notation in this style of CG: In the most basic combinatoric mode, (1) if α is an expression of type X/Y , and β is an expression of type Y , then $\alpha\beta$ is an expression of type X and (2) if α is an expression of type $X \backslash Y$, and β is an expression of type Y , then $\beta\alpha$ is an expression of type X .



Here too, because of modification and the composition employed, e.g. in Steedman’s (1996, 2000) analysis of filler-gap phenomena, subcategorization dependencies are extended over an unbounded domain in predictable ways.

In sum, I assume that grammatical theory must include some hypothesis about the domain in which subcategorization dependencies hold. Any such hypothesis involves basic lexical subcategorization restrictions which function within narrowly specified domains and which interact with other grammatical principles to account for the fact that local subcategorization domains are extended in precisely characterizable ways to allow extensions of local constraints in coordinate structures, modification structures, and so forth. The mechanisms for handling basic subcategorization dependencies vary from theory to theory, ranging from the pristine lexical categories of CG to the arcane theory of lexical insertion presented by Chomsky 1965), which is formulated in terms of pre-terminal phrase markers, lexical substitution transformations, and a matching condition. The mechanisms for extending local subcategorization domains also differs from theory to theory: the work is done variously by transformations (Chomsky/Kajita’s model), adjunction (TAGs), composition and modification (CG) and general principles of feature inheritance (GPSG). HPSG analyses (e.g. that of Pollard and Sag (1994)) have attempted to integrate the basic lexical subcategorization mechanism of CG (reformulated slightly in terms of valence lists) with the general principles of feature inheritance that were pioneered within GPSG.

As we will see, the locality of basic subcategorization restrictions, and with it the locality of agreement, case assignment, and government, raises a variety of issues. In essence, the fundamental locality of these phenomena follows from the nature of the arguments on valence lists, as in CG. I will refine this idea in the proposal that follows directly.

1.3 Locality of Construction

Since the inception of work in HPSG, it has been assumed that there are two kinds of signs – words and phrases, with the feature DAUGHTERS (DTRS) being appropriate only for the type *phrase*. Grammar schemata were introduced in Pollard and Sag 1994 as the HPSG analog of grammar rules. These schemata specified an inventory of phrase types, where phrases had the geometry shown in Figure 1.8. Since (1) phrases contained daughter structures of arbitrary depth and (2) schemata imposed constraints directly on phrases, there was nothing in this set-up that imposed any notion of locality. Putting this in more familiar terms, in this ‘standard’ version of HPSG, one could write grammar rules like those in (15), where the right-hand side of the rule needn’t be confined to a sequence of categories (as in CFG), but could in fact be a structure of arbitrary complexity:

- (15) a. $VP \rightarrow V [{}_S NP_{sing} VP]$
 b. $VP \rightarrow V [{}_{NP} Det Adj N]$
 c. $S \rightarrow NP [{}_{VP} V [{}_S [{}_{NP} NP Conj NP] VP]]$

Nothing but an unspoken ‘gentleman’s agreement’ prevented the HPSG grammarian from writing a schema that directly referenced a daughter’s daughters, or in fact elements that appear at any arbitrary depth of embedding.

HPSG had thus evolved far from its GPSG (CFG) roots, an evolutionary path that did not go unnoticed. For example, Copestake (1992) observed that:

[...] it is unclear that the HPSG account of phrasal signs as feature structures which incorporate their daughters is the best one to adopt. Constraint resolution can be used to perform operations which cannot be straightforwardly mimicked by more conventional grammar rules. [...]. However, it is not clear to me whether HPSG currently takes advantage of this possibility in any very significant way. There have to be good reasons to adopt an approach which makes most known parsing technology inapplicable.

Copestake’s observation still has force today, though of course there is now considerable work developing analyses based on linearization theory,⁷

⁷ See, for example, Reape 1994, 1996, Kathol 2000, and Daniels and Meurers 2004.

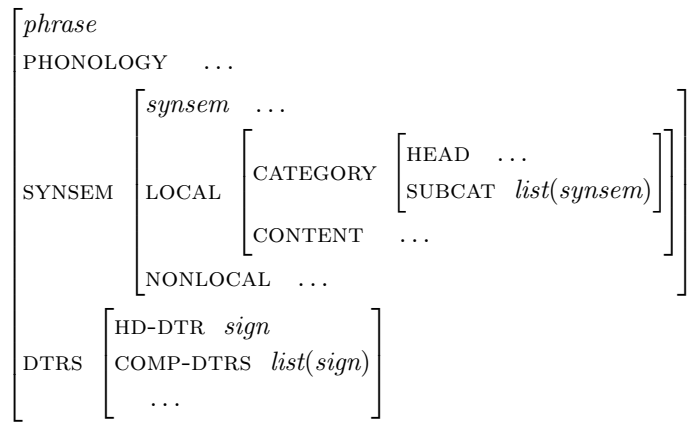


FIG. 1.8 Feature Geometry of Pollard and Sag 1994

which uses a DOMAIN feature to allow ‘liberation’ of embedded elements, making them locally accessible at ‘higher’ levels of tectogrammatical derivation.⁸ Apart from this line of research, there are to my knowledge no HPSG analyses that propose a grammatical schema making direct reference to embedded structure. The practice of the HPSG community seems to adhere to the notion of locality that is inherent in CFGs. English tag questions pose an interesting challenge to constructional locality, since they involve agreement between the main clause subject and the subject pronoun realized within the tag:

- (16) a. He is going to get into trouble, isn’t he/*she/*it?
 b.*He is going to get into trouble, aren’t they/you/we?

Bender and Flickinger (1999) assume that the agreement between the two subjects is syntactic, and hence that the two verbs and the two subjects in any tag question must all agree. This view, however, is inconsistent with well known data like (17), which argues that the agreement in question is semantic, rather than syntactic:⁹

- (17) a. Sears is open, aren’t they?
 b. At least one of us is sure to win, aren’t we?

⁸ For critical discussion of this approach, see Müller 2004, 2005.

⁹ See Kay 2002 and the references cited there.

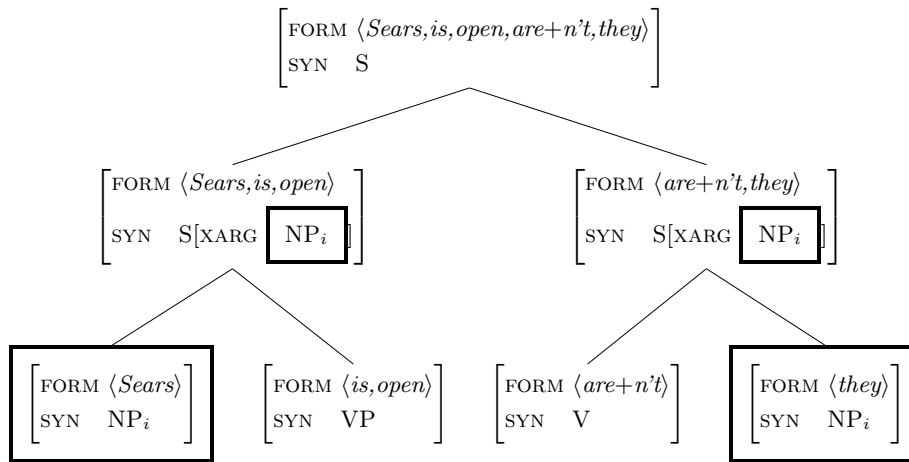


FIG. 1.9 A Tag-Question

But however the agreement in question is to be analyzed, the agreement relation between the two subjects is non-local, i.e. it involves agreement between the two boxed NP constituents shown in Figure 1.9.

As Bender and Flickinger argue, the English tag-question construction is evidence not for an analysis in terms of nonlocal constraints, but rather for a treatment in terms of a feature that ‘passes up’ information about the subject NP to the clausal level, indicated here via the boxed values of the feature XARG, discussed further below. Under such an analysis it is possible to treat the agreement in tag questions locally, i.e. via a local constraint requiring the relevant identity (coindexing) between the XARG value of the main clause and that of the tag clause (the two daughters of the root S in Figure 1.9).

1.4 Sign-Based Construction Grammar

Here I sketch a version of grammatical theory, building on the modeling assumptions developed within the HPSG research community, that incorporates a strong version of locality. To this end, phrases are distinguished from the structures associated with them. Phrases, like words, are signs and hence specify values for the features PHONOLOGY,

$$\left[\begin{array}{ll} \textit{sign} & \\ \text{PHONOLOGY} & \textit{phonological-phrase} \\ \text{FORM} & \textit{list(morph-form)} \\ \text{SYNTAX} & \textit{syn-obj} \\ \text{SEMANTICS} & \textit{sem-obj} \\ \text{CONTEXT} & \textit{context} \end{array} \right]$$

FIG. 1.10 The Sign in SBCG

FORM, SYNTAX, SEMANTICS, and CONTEXT, but, crucially, not DAUGHTERS (and herein lies the key departure from previous work in HPSG). A construction, like a schema in Pollard and Sag 1994, is a constraint licensing a local pattern of sign combination. That is, a construction places restrictions on what properties signs must have if they are to directly combine with one another (or, to use minimalist terminology, to ‘externally merge’). A construction may in addition place constraints on the sign that results from such a combination. On this conception of grammar, a construction is a CFG-like grammar rule that provides a particular set of constraints on the form, syntactic category, meaning, and use conditions of the mother sign, stated in terms of the properties of its daughters. The objects defined by constructions are thus configurations of signs: a set of daughter signs and one more sign that is the mother of those daughters. Let us call each such configuration a ‘construct’.

Once this distinction is recognized, it becomes possible to adopt a simpler feature geometry like the one proposed in Pollard and Sag 1987, eliminating the feature SYNSEM. In fact, I will eliminate (following Sag 2007) a number of other features that have appeared in HPSG analyses, including LOCAL, NONLOCAL, and HEAD. The resulting feature geometry I assume here is sketched in Figure 1.10. In addition, constructs will be modeled as feature structures, as shown in Figure 1.11.¹⁰ This last move is easily achieved by the type declarations sketched in Figure 1.12.

A few words of explanation for readers not familiar with these notions and notations: I assume a grammar contains a ‘signature’ (like the key and time signatures in a piece of music) that spells out the general nature of the objects in the language model, assumed here to be a collection of feature structures each of which is either an atom (e.g. *accusative*, +) or

¹⁰ For expositional purposes, I will sometimes represent constructs in tree notation and will use SYNTAX and SEMANTICS values, as in Figure 1.11.

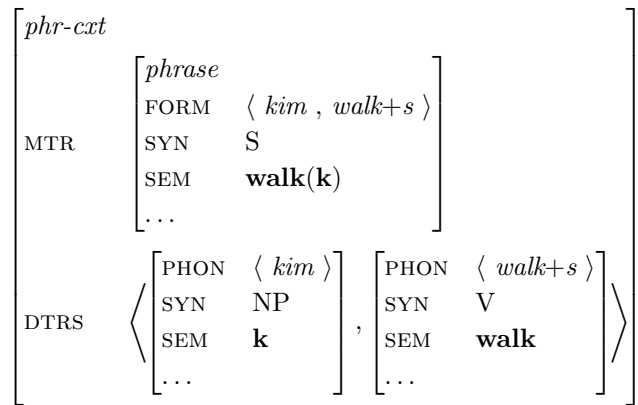


FIG. 1.11 A Clausal Construct

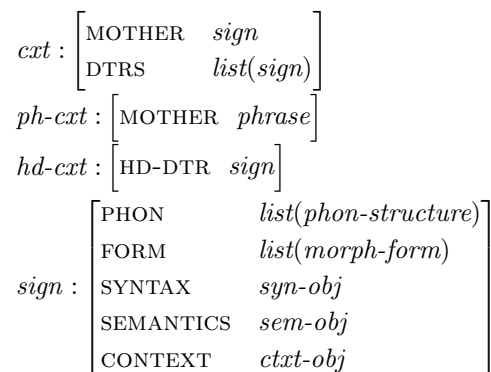


FIG. 1.12 Type Declarations

a function (e.g. the phrasal construct sketched in Figure 1.11). This work is done in terms of a set of types and a ‘declaration’ (for each of the nonatomic types) specifying which features are appropriate for feature structures of that type and what kind of value each of those features is mapped to. Thus, the last specification in Figure 1.12 declares (1) that there is a type of feature structure called *sign*, (2) that the domain of functions of this type includes the features PHONOLOGY, FORM, SYNTAX, SEMANTICS, and CONTEXT, and (3) that any value assigned to each of these features must be of the indicated type. *list(morph-form)* indicates (that any value assigned to the feature FORM must be) a (possibly empty) list, each of whose members must be of type *morph-form*.

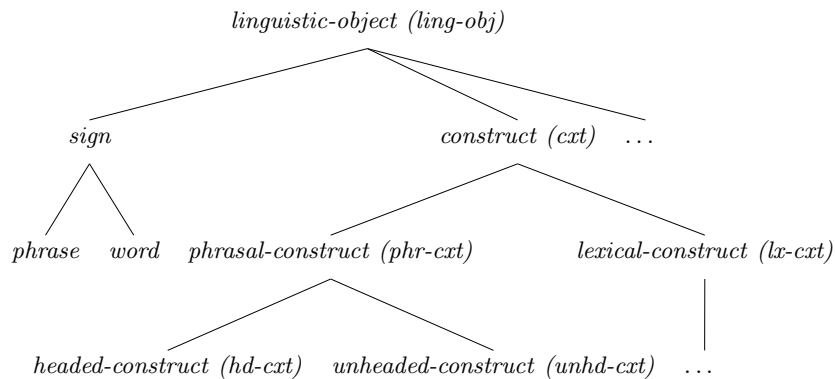


FIG. 1.13 A SBCG Type Hierarchy

The diagram in Figure 1.13 specifies a hierarchical classification of the types which interacts with the type declarations in Figure 1.12. A feature structure assigned to a type T also exhibits the properties that the signature prescribes for the supertypes of T . Thus, words have properties of their own, but they also exhibit the properties of signs; headed constructs must also obey the properties of both phrasal constructs and constructs in general, and so forth.

A SBCG defines complex expressions including a principle like (18), which allows recursive application of constructions:

(18) The Sign Principle:

Every sign must be lexically or constructionally licensed, where:
 a sign is lexically licensed only if it satisfies some lexical entry and
 a sign is constructionally licensed only if it is the mother of some construct.

This framework has come to be known as **Sign-Based Construction Grammar (SBCG)**,¹¹ though of course it is still a kind of HPSG, given that it embodies signs, linguistically motivated types, type constraints, and a hierarchically organized lexicon, *inter alia*.

¹¹ For an early formulation, see Chapter 16 of Sag, Wasow, and Bender 2003, which develops ideas first presented in Sag 2001. For a more detailed exposition, see Sag 2007.

It follows from SBCG, as a matter of principle, that a construction cannot have direct access to properties of a mother and its granddaughters. If we observe that there is some such dependency, then we must provide an analysis in terms of some property of the granddaughter that is systematically encoded on the daughter, and hence rendered locally accessible at the higher level. This has the virtue of making explicit exactly where nonlocality resides in grammatical structures. It also fosters the development of general principles constraining the distribution of feature specifications across constructs. In fact, the fundamental principles of Pollard and Sag 1994 are now recast as constraints on constructs, as shown in (19):¹²

(19) a. **Head Feature Principle:**

$$hd\text{-}cxt \Rightarrow \left[\begin{array}{l} \text{MTR} \quad \left[\text{SYN|CAT} \quad \boxed{\mathbb{1}} \right] \\ \text{HD-DTR} \quad \left[\text{SYN|CAT} \quad \boxed{\mathbb{1}} \right] \end{array} \right]$$

b. **Subcategorization Principle:**

$$hd\text{-}cxt \Rightarrow \left[\begin{array}{l} \text{MTR} \quad \left[\text{SYN|VAL} \quad \boxed{\mathbb{A}} \right] \\ \text{DTRS} \quad \boxed{\mathbb{E}} \circ \langle \mathbb{1} \rangle \\ \text{HD-DTR} \quad \boxed{\mathbb{1}} \left[\text{SYN|VAL} \quad \boxed{\mathbb{A}} \oplus \boxed{\mathbb{E}} \right] \end{array} \right]$$

The effect is the same as in PS-94: the head daughter projects certain featural information (the HEAD feature specifications) to its mother and the VALENCE list of the head daughter is matched against the other daughters, with the remaining members of that list being passed up to be the mother’s VALENCE list.

Finally, SBCG also provides a precise way of formulating lexical rules, i.e. by treating them as varieties of lexical construction.¹³ We may posit three subtypes of lexical construct: *inflectional-construct*, *derivational-construct*, and *post-inflectional-construct*, each with its own properties. Following in the main Sag et al. 2003 (see especially Chapter 16), we may assume that lexical entries in general describe feature structures of type *lexeme* (rather than *word*). Hence derivational constructions involve constructs (of type *deriv-cxt*) whose mother is of

¹² Note that the feature SUBCAT is replaced by VALENCE (VAL). ‘ \circ ’ is Reape’s domain union operator: ‘ $\boxed{\mathbb{A}} \circ \boxed{\mathbb{E}}$ ’ is satisfied by any list containing exactly the elements of $\boxed{\mathbb{A}}$ and $\boxed{\mathbb{E}}$, as long as any α which precedes some β in $\boxed{\mathbb{A}}$ or in $\boxed{\mathbb{E}}$ also precedes β in $\boxed{\mathbb{A}} \circ \boxed{\mathbb{E}}$. ‘ \circ ’ is thus a ‘shuffle’ operator.

¹³ See Copestake 1992 and Meurers 1999.

type *lexeme*; inflectional constructions involve unary constructs (of type *infl-ctx*) whose mother is of type *word* and whose daughter is of type *lexeme*; and post-inflectional constructions involve unary constructs (of type *post-infl-ctx*) where both mother and daughter are of type *word*. This proposal thus provides a unified approach to the construction of words and phrases, allowing for hierarchical generalizations of varying grain, without the need for ancillary devices.

1.5 Predictions of Locality

The syntactic objects of SBCG (modeled as feature structures of type *syn-obj*) are the values of the feature SYN. These feature structures include specifications for category and valence information, as illustrated in (20):

$$(20) \quad \left[\begin{array}{c} \textit{sign} \\ \text{SYN} \left[\begin{array}{cc} \textit{syn-obj} \\ \text{CAT} & \textit{syn-cat} \\ \text{VAL} & \textit{list}(\textit{sign}) \\ \dots \end{array} \right] \end{array} \right]$$

Category and valence information are within the sign and lexical subcategorization obeys the Subcategorization Principle in (19b) above, which identifies the head’s VAL specifications with the signs (not the constituent structure) of the selected elements. It therefore follows that a complement’s category and valence information is accessible to a subcategorizing head, and that information associated only with elements used to construct that complement is not. This circumscription of information is quite like CFG, where a rule like (21) has no access to information about which of the rules in (22) will be used to expand the complement \bar{N} :

$$(21) \quad \bar{V} \rightarrow V^0 \bar{N}$$

$$(22) \quad \begin{array}{l} \bar{N} \rightarrow \text{Det } \bar{N} \\ \bar{N} \rightarrow \bar{N}^+ C \bar{N} \\ \bar{N} \rightarrow \bar{N} \bar{P} \\ \dots \end{array}$$

SBCG thus embodies a strong theory of the locality of category selection in the normal sense of that term. The predictions of course rely crucially on the HFP as well, in much the same way as earlier work in GPSG and HPSG. That is, among SBCG’s HEAD features are CASE, VFORM, PFORM, PRED, and AUX, and a phrase’s value for HEAD must be the same as that of its head daughter in order for the HFP to be satisfied. For example, verbs like *depend* or *rely* require that the prepositional head within their PP complement be *on* or *upon* and this is ensured by a lexical specification like the one in (23):

$$(23) \quad \left[\begin{array}{l} \text{FORM} \quad \langle \textit{rely} \rangle \\ \text{SYN} \quad \left[\text{VAL} \quad \langle \text{NP} , \text{PP}[\text{PFORM } \textit{on}] \rangle \right] \end{array} \right]$$

Similarly, the modal verbs select for a VP complement whose verbal head is specified as [VF *base*]. This will have the intended effect on the VP’s head daughter, as sketched in Figure 1.14. Note in addition that since a verb’s VAL list includes reference to its subject (the first valent), the domain of locality is automatically extended to include subjects without the introduction of the AGR feature discussed earlier in connection with GPSG.

To see how this set-up also imposes locality on agreement, consider the following well attested agreement patterns:

- (24) a. Verb-subject agreement
- b. Verb-object agreement
- c. Noun-possessor agreement
- d. Determiner-noun agreement
- e. Modifier-modified agreement

Following the long-term practice of the constraint-based grammar community, including LFG, GPSG/HPSG, TAG, and CCG, among others, all these phenomena have been analyzed in terms of feature compatibility. In the GPSG/HPSG tradition, the particular method of analysis has involved features of selection (AGR, SUBCAT, VAL, MOD, SPEC, SELECT, etc., depending on the particular proposal). In all such analyses, the selecting sign specifies a value for one of these features that is identified with the relevant part of the selected element. For example, a third-singular verb (e.g. *runs*) is specified as in (25) and the Subcategorization Principle requires that the VAL value of *runs* be identified with the sign of its subject:

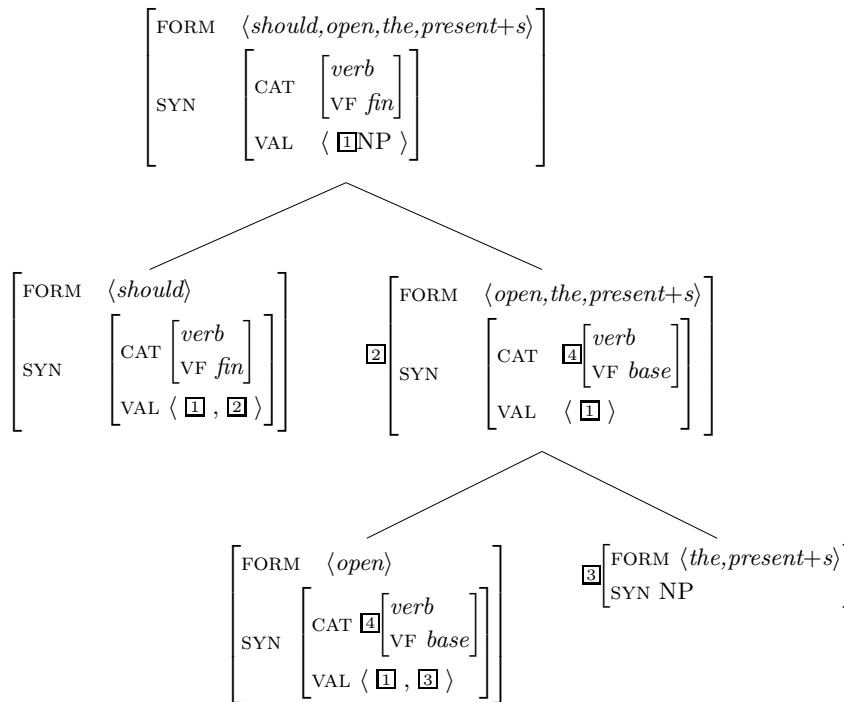


FIG. 1.14 Interaction of the Subcategorization and Head Feature Principles

(25)
$$\left[\begin{array}{l} \text{FORM} \langle \textit{run+s} \rangle \\ \text{SYN} \left[\begin{array}{l} \text{VAL} \langle \text{NP}[\textit{3rd, sing}] \rangle \end{array} \right] \end{array} \right]$$

Subject-verb agreement is thus treated by the very same mechanisms as case government and category selection – that is via simple specifications in the lexical entries of agreeing elements, governors, or modifiers, as described above. No special ‘agreement theory’ needs to be introduced for verb-subject agreement, verb-object agreement, or indeed agreement between a head and any of its valents.

Notice that this analysis also embodies a clear notion of directionality. The subject NP (in English) bears certain feature specifications ‘inherently’, while a verb that agrees with the subject NP specifies its

requirements in terms of the ‘selection’ feature VAL. This parallels the intuitive directionality of government: a verb that requires a quirky subject or object case uses VAL to specify those requirements, while the governed valent bears its case specification inherently.

Let us now turn to other kinds of agreement. I follow the economical and insightful analysis of Van Eynde, who employs the non-valence feature SELECT.¹⁴ SELECT is used to let an expression select what it can modify or combine with as a ‘marker’. The SELECT value of a modifier, a specifier, or a marker is a sign and this value must be identified with the head daughter in a ‘head-functor’ construct. Given this analysis, agreement is again accounted for in terms of lexical entries for the agreeing elements that use SELECT to restrict the range of the elements to be modified, specified, or simply ‘marked’, as illustrated in (26):¹⁵

- (26) a. French feminine plural adjective:
- $$\left[\begin{array}{l} \text{FORM} \quad \langle \textit{grand}+e+s \rangle \\ \text{SYN} \quad \left[\text{CAT} \quad \left[\text{SELECT} \quad \text{N}[\textit{fem},\textit{pl}] \right] \right] \right]$$
- b. English marker:
- $$\left[\begin{array}{l} \text{FORM} \quad \langle \textit{that} \rangle \\ \text{SYN} \quad \left[\text{CAT} \quad \left[\text{SELECT} \quad \text{S}[\textit{fn}] \right] \right] \right]$$
- c. English determiner:

¹⁴ Following Van Eynde (1998), who builds directly on Allegranza 1998, the features MOD and SPR of Pollard and Sag 1994 are replaced by the single feature SELECT (SEL). See also Van Eynde 2006, 2007 and Allegranza 2007.

¹⁵ Pullum and Zwicky (1988) suggest that the *a/an* alternation should not be analyzed via a condition that is part of the determiner’s lexical entry. They suggest that such a condition should be impossible because it would refer to the following syntactic context. Instead, they offer a condition on shape that overrides the lexical entry for the indefinite article and stipulates that another shape is called for.

However, Pullum and Zwicky offer no argument against a lexical *a/an* analysis such as the one presented here, which provides a straightforward account of the relevant data, including such contrasts as *an interesting suggestion* vs. *a clever idea* (since the SELECT value corresponds to the entire CNP that the determiner combines with syntactically). Moreover, the lexical analysis proposed here avoids any appeal to competition among alternative forms, ‘overriding’ or other nonmonotonic devices.

$$\left[\begin{array}{l} \text{FORM} \langle a \rangle \\ \text{SYN} \left[\begin{array}{l} \text{CAT} \left[\begin{array}{l} \text{SELECT} \left[\begin{array}{l} \text{PHON}[\text{FIRST } \textit{consonant}] \\ \text{SYN} \left[\begin{array}{l} \text{CAT} \left[\begin{array}{l} \textit{noun} \\ \text{MRKG } \textit{unmkd} \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right]$$

Assuming that agreement information is specified in terms of category features like PERS, NUM and GEND, all such agreement is localized through the interaction of the Subcategorization Principle and the Head Feature Principle.¹⁶ Crucially, however, this means that agreement phenomena, like government and category selection, are local in their basic case, but are extended within complex structures by other, independently motivated grammatical principles to induce indirect, long distance agreement. Notable among such examples is the long distance agreement of reflexive pronouns that is mediated by the theories of binding, control and subcategorization in examples like (27):

- (27) $\left\{ \begin{array}{l} \text{They} \\ *She \\ *He \end{array} \right\}$ may want to consider trying to get **themselves** on the ballot.

Finally, let us consider the issue of semantic selection, which intuitively exhibits a locality constraint as well. For example, intuition tells us that there is no verb in any human language that requires its second argument to be a proposition built up from a relation whose second argument is itself a proposition. That is, we assume that no human language could be just like English except for the inclusion of a verb *hink* whose selection properties give rise to semantic contrasts like the following:

- (28) a. Kim hinks that Sandy believes that the earth is flat.
 b. #Kim hinks that Sandy died.
 c. #Kim hinks that Sandy loves Pat.

And it seems equally unlikely that there could be an English-like language that included a verb *fask* which, though similar to *ask*, would require its second argument to be an animate *wh*-question, determining semantic contrasts like the following:

¹⁶ If, on the other hand, some agreement information is encoded via semantic indices (as proposed by Pollard and Sag (1994)), then something more needs to be said to ensure locality of agreement. See below.

- (29) a. Bo fasked who left.
b.#Bo fasked whether Carrie had left.
c.#Bo fasked what Carrie had left.

Though we should not lose sight of the fact that these are only intuitions, the intuitions are nonetheless quite robust. And if semantic composition proceeded entirely in terms of the unstructured senses of the immediate constituents of an expression (as in Montague Grammar and its various descendants), then SBCG would indeed predict that semantic selection obeyed the same locality constraints as the phenomena just discussed. However, if compositionally derived meanings are more structured, as is often assumed in order to solve the problem of individuating propositions at a sufficiently fine grain, then the meanings of the constituents of an expression might well be visible (albeit hard to identify) to a selecting element.¹⁷ Structured meanings thus have the potential to vitiate the predictions of SBCG with respect to the locality of semantic selection. I regret being able to do no more at present than to flag this issue, hoping that it will be clarified by future research.

1.6 Some Analytic Challenges

The theory of SBCG is attractive for its simplicity, precision, and predictive power, yet there are various empirical phenomena that, at least in their outward appearance, appear to defy the localism embodied in SBCG. In the remainder of this chapter, I will examine a number of such phenomena, showing that an attractive localist analysis within SBCG is readily available.

1.6.1 *Nonlocal Case Assignment in English*

English *for/to* clauses present an interesting analytic challenge for the locality of case assignment. In order to analyze contrasts like the one in

¹⁷ Within the framework of Minimal Recursion Semantics (MRS – see Copestake, Flickinger et al. 2005 and Copestake, Lascarides et al. 2002), all the semantic predications of an embedded phrase are present on its RELATIONS list, and hence are ‘locally visible’, just as they are in earlier approaches to semantics within HPSG. Particular levels of embedding would be quite difficult to identify within MRS, where the order of elements on the RELATIONS list has no semantic significance. Hence one might argue that MRS provides sufficient prediction in this domain.

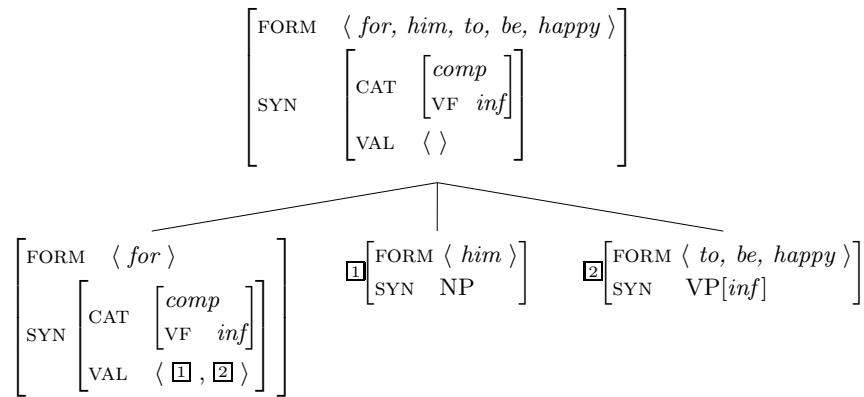


FIG. 1.15 A *For-To* Clause

(30), it is necessary that an accusative case constraint be imposed somehow:

- (30) a. I prefer [for [*they to be happy]]
 b. I prefer [for [them to be happy]].

But given the standardly assumed structure in (30), the subject NP of the infinitive is not locally accessible to the complementizer *for*, which selects for the infinitival S either as a head (via VAL) or as a marker (via SPEC). Nor can the infinitive marker *to* assign accusative case to its subject, as in examples like (31), that subject must be compatible with nominative case:

- (31) [He/*Him seems [to be happy]].

Sag (1997) argues that the standard structure for *for/to* clauses should be replaced by the flat head-complement structure in Figure 1.15.¹⁸ Assuming this structure, rather than the one in (30), the lexical entry for the complementizer *for* can simply require that its first VALENCE element be an accusative NP. The problematic NP is now locally accessible.

¹⁸ Here and throughout this section, I have regularized valence features and the attendant feature geometry to conform with the preceding discussion.

Moreover, the structure in Figure 1.15 is independently motivated, for it provides an immediate account of contrasts like the following, first noted by Emonds (1976):

- (32) a. Mary asked me [if, in *St. Louis*, [John could rent a house cheap]].
 b. He doesn't intend [that, in these circumstances, [we be rehired]].
 c.*Mary arranged for, *in St. Louis*, John to rent a house cheap.
 d.*He doesn't intend for, *in these circumstances*, us to be rehired.

Assuming that only finite CPs have the traditional structure indicated in (32a-b), there is no constituent for the italicized modifiers to modify in (32c-d). The deviance of these examples follows from the same constraints that disallow the indicated modifiers in (33a-b), whose structure is analogous to the new *for/to*-clausal structure:

- (33) a.*Kim persuaded *in St. Louis* Sandy to rent a house cheap.
 b.*Lee believed *in these circumstances* Sandy to be in the right.

1.6.2 The Local Registration of Filler-Gap Dependencies

Over the last thirty years, it has been shown that numerous languages exhibit phenomena that are sensitive to the presence or absence of an extraction path – the part of the grammatical structure connecting the filler and the ‘gap’ in a filler-gap dependency. These phenomena include:

- (34) Irish complementizer selection (McCloskey 1979, 1990)
 French ‘stylistic’ inversion (Kayne and Pollock 1978).
 Spanish ‘stylistic’ inversion (Torrego 1984)
 Kikuyu downstep suppression (Clements 1984, Zaenen 1983)
 Chamorro verb agreement (Chung 1982, 1998¹⁹)
 Yiddish verb-subject inversion (Diesing 1990)
 Icelandic expletive constructions (Zaenen 1983)
 Adyghe ‘*wh*-agreement’ (Polinsky 2007)

It has often been pointed out (see Zaenen 1983, Hukari and Levine 1995, Bouma et al. 2001, and Levine and Hukari 2006), that constraint-based

¹⁹ But see Donohue 2003 and the references cited there for a critical assessment of Chung’s analysis.

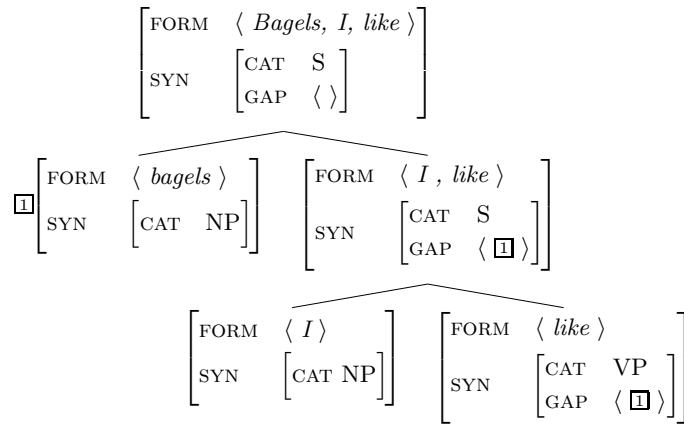


FIG. 1.16 A Filler-Gap Dependency in SBCG

accounts of extraction provide a straightforward account of phenomena sensitive to extraction paths without the introduction of otherwise unmotivated intermediate traces. For example, in GPSG/HPSG/SBCG analyses, the information about a FG dependency is locally encoded along the extraction path, as shown in Figure 1.16.²⁰ This provides a set of local syntactic distinctions that are suitable for analyzing the critical effect of extraction path domains on lexical choice and constructional options.

Under these assumptions, the Irish complementizers are distinguished in terms of sentential complements – S[GAP ⟨X⟩] vs. S[GAP ⟨⟩], for example. And a lexical entry for a gap-binding predicate like English *easy* can be formulated that requires its complement to contain a NP-type gap, i.e. to be specified as [GAP ⟨NP⟩]. Note, however, that it is not possible to write a lexical entry that requires a gap appearing at some fixed level of embedding. That is, the ‘localist’ analysis of filler-gap dependencies that has emerged from the GPSG/HPSG tradition comes close to predicting (correctly, to the best of my knowledge) that no grammar for a natural language can impose an arbitrary depth on a filler-gap dependency. The positions in which the gap can appear are

²⁰ This figure illustrates a traceless analysis of English topicalization. Analyses that include *wh*-traces, e.g. those of Pollard and Sag 1994 or Levine and Hukari 2006, are also wholly consistent with SBCG.

always determined by general constraints on the ‘inheritance’ of GAP (or SLASH) specifications.²¹

1.6.3 Case Stacking Languages

One of the best known examples of apparent nonlocal case assignment is the phenomenon of case ‘stacking’, as in the following examples from Martuthunira, a Pama-Nyungan language:²²

(35)

Ngayu nhuwa-lalha tharnta-a kupuyu-marta-a thara-ngka-marta-a.
 1SG.NOM spear-PAST euro-ACC little-PROP-ACC pouch-LOC-PROP-ACC
 ‘I speared a euro with a little one in its pouch.’
 (Dench and Evans 1988)

(36)

Ngunhu wartirra puni-lha ngurnu-ngara-mulyarra kanyara-ngara-mulyarra
 the woman go-PAST that-PL-ALL man-PL-ALL
 kapunmarnu-marta-ngara-mulyarra jirli-wirra-marta-ngara-mulyarra.
 shirt-PROP-PL-ALL arm-PRIV-PROP-PL-ALL
 ‘That woman went towards those men with shirts without sleeves.’
 (Andrews 1996)

The operant generalization about these examples is that nominals within NPs are inflected not only in accordance with their local grammatical function, but also so as to reflect the function of the NPs that contain them. The unbounded case dependency phenomenon illustrated in (35)–(36) seems to pose a serious challenge for any locality hypothesis, including the one entailed by the interaction of the Subcategorization Principle and the HFP.

However, an elegant analysis of this phenomenon in terms of purely local constraints has been developed by Malouf (2000). Malouf proposes that in case stacking languages the value of the feature CASE is not an atomic case, but rather a list of such atoms. Assuming that nouns select for their NP dependents, the lexical entry for the noun *tharnt* ‘euro’ looks like (37):

²¹ This should be compared with a different approach that could also be incorporated within HPSG, namely the use of regular expressions to characterize the relation between fillers and gaps. Under this alternative (cf. its deployment within LFG under the rubric of ‘functional uncertainty’), no such prediction is made, as one could write a lexical entry that forced that gap to appear at some fixed depth within the infinitival complement of *hard*.

²² A euro is a kind of marsupial distinct from kangaroos, wallabies, pademelons, and potoroos.

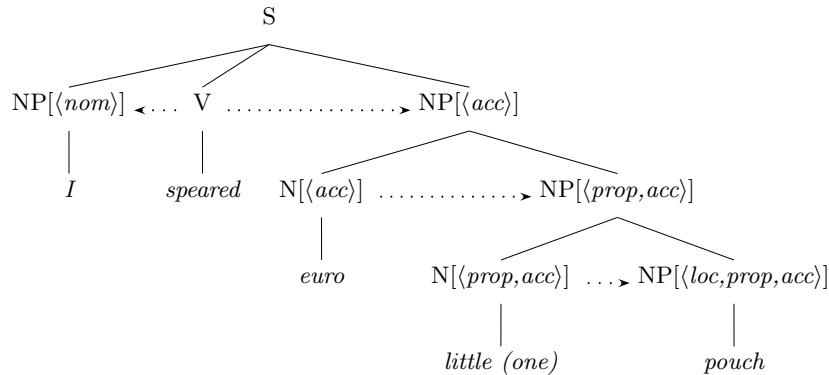


FIG. 1.17 Case Government in Martuthunira

$$(37) \left[\begin{array}{l} \text{PHON} \langle \text{tharnt-} \rangle \\ \text{SYN} \left[\begin{array}{l} \text{CAT} \left[\begin{array}{l} \textit{noun} \\ \text{CASE} \boxed{E} \end{array} \right] \\ \text{VAL} \langle \text{NP}[\text{CASE} \langle \textit{prop} \rangle \oplus \boxed{E}] \rangle \end{array} \right] \end{array} \right]$$

The key thing to see here is that every word formed from this stem will bear a particular case specification that is then passed on to the NP on that word’s VAL list.

Malouf’s treatment of nouns interacts with the analysis of verbs, which is sketched in (38):

$$(38) \left[\begin{array}{l} \text{PHON} \langle \textit{nhuwalalha} \rangle \\ \text{SYN} \left[\begin{array}{l} \text{CAT} \left[\begin{array}{l} \textit{verb} \\ \text{CASE} \boxed{E}(\) \end{array} \right] \\ \text{VAL} \langle \text{NP}[\langle \textit{nom} \rangle \oplus \boxed{E}], \text{NP}[\langle \textit{acc} \rangle \oplus \boxed{E}] \rangle \end{array} \right] \end{array} \right]$$

Finite verbs bear an empty CASE specification. However, (38) is formulated so as to illustrate the general principle that lexical heads add their own CASE value to that of their dependents. As a result of this case addition, CASE values become longer with embedding, as shown in Figure 1.17.

Long-distance case stacking is thus a consequence of CASE specifications that pass the case properties of a superordinate context down into a subordinate one, adding only the case information that reflects the local grammatical function of a given head-dependent combination. The

morphological case inflections are based on local CASE specifications, just as they are in languages that lack case stacking. But when multiple case affixes are present (e.g. on *pouch* in Figure 1.17), it follows that the CASE specification of the noun is non-singleton. This in turn entails that the immediately embedding syntactic context (e.g. *little (one)*) must introduce an appropriate case specification. Otherwise, the maximal NP in (1.17) would fail to meet the VALENCE requirements of the verb *spearred*. The local constraints of lexical items and general grammatical principles thus interact to guarantee a long-distance case dependency that is bounded only by the complexity of the embedding environment.

1.6.4 *The Role of Subjects*

Earlier, I mentioned the presumed locality of semantic role assignment. However, as a number of researchers have recently argued, there are phenomena in a variety of languages whose analysis requires, for example, that a verb selecting a sentential complement must be able to place constraints on the subject realized within that complement. One of these is English ‘copy raising’ (Rogers 1974, Potsdam and Runner 2001, Asudeh 2002), illustrated in (39):

- (39) There looks like there’s going to be a storm/*it’s going to rain/*Kim’s going to win.

Also relevant are controlled pronominal subjects in Serbo-Croatian (Zec 1987), Halkomelem Salish (Gerdtts and Hukari 2001) and other languages, where a control verb requires that the subject pronoun realized within its clausal complement be coindexed with one of the other arguments of the control verb (its subject (*promise*-type) or its object (*persuade*-type)), as shown in (40):

- (40) a. NP_i promise [Comp he_i VP]
 b. NP persuade NP_i [Comp he_i VP]

The problems of raising across Polish prepositions (Przepiórkowski 1999, Dickinson 2004), and complementizer agreement in Eastern Dutch dialects (Höhle 1997) are similar: a particular argument realized within a given expression must be ‘visible’ to an external entity that combines with that expression. Moreover, as is well known, there are many English idioms that require referential and agreement identity between a subject and a possessor within an object NP, or which assign a semantic role to the object’s possessor. These are illustrated in (41):

- (41) a. He_i lost [his_i/*her_j marbles].

- b. They_i kept/lost [their_i/*our_j cool].

A principled solution to all of these problems, suggested independently by a number of these researchers, is the introduction of a feature (distinct from VAL) that passes up to a given phrase information about one of the daughters used to construct that phrase. Kiss (1995) proposed such a feature for the subject of nonfinite verbal clauses in German, calling it SUBJECT, and this feature has been used by Meurers (1999, 2001) and others to make subjects accessible at higher levels of structure.²³

However, it would be desirable to use the same feature to make genitive pronouns realized within a given NP available for selection by elements outside that NP. In addition, the Polish preposition raising phenomenon discussed by Przepiórkowski (1999) and Dickinson (2004) motivates an analysis where the object of certain prepositions is available for selection by elements external to the PP that the preposition projects. In sum, there is some variation as to which element within a phrase is externally accessible. Since ‘subject’ is too narrow a notion empirically, SUBJECT is an inappropriate name for the feature in question. I have previously proposed instead to name the relevant feature EXTERNAL ARGUMENT (XARG).²⁴ Because XARG is a category feature, it percolates information about a designated phrasal constituent, as illustrated in Figure 1.18.

Assuming, following Pollard and Sag (1994), that there are three subtypes of the type *index (ref (referential-index), it (expletive-it-index), and there (expletive-there-index))*, the copy raising examples mentioned in (39) above can be treated simply by associating the relevant lexical entry for *looks (like)* with the VAL list in (42):

$$(42) \quad \left[\text{VAL} \left\langle \text{NP}_i, \left[\begin{array}{c} \text{S} \\ \text{XARG NP}[\text{pro}]_i \end{array} \right] \right\rangle \right]$$

And if an object NP includes information about its (prenominal) possessor in its XARG value, then an idiomatic verb like *lose* can be specified as in (43):

²³ Kiss’s proposal is an extension of earlier proposals that have been made within GPSG/HPSG, e.g. the AGR feature used by Gazdar et al. (1985) and Pollard’s (1994) ERG feature.

²⁴ Sag and Pollard (1991) proposed a semantic feature EXTERNAL-ARGUMENT (XARG), which made only the index of the subject argument available at the clausal level. This analysis has been incorporated into Minimal Recursion Semantics (and the English Resource Grammar) by Flickinger and Bender (2003).

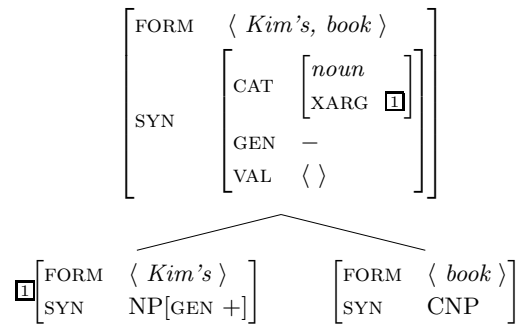


FIG. 1.18 XARG Analysis of Genitive-Embedding NP

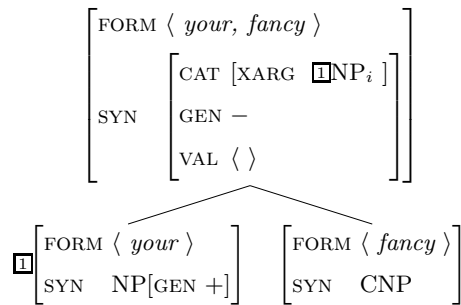
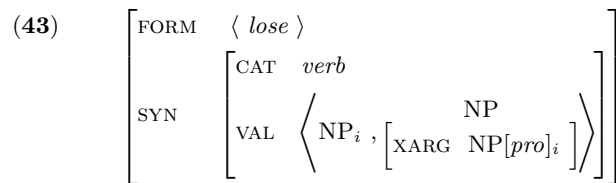


FIG. 1.19 XARG Analysis of *your fancy*



Similarly, an idiomatic verb like *tickle* can assign a semantic role to its object’s possessor. In both cases, all that is required is that the NP’s XARG value be identified with the NP’s possessor, as sketched in Figure 1.19.

All of the phenomena just enumerated, in addition to the English tag-question construction discussed earlier, provide motivation for XARG specifications as part of the CAT value of sentential and NP signs. Note that the XARG value (either a sign or the distinguished atom *none*) differs from the VAL value (a list of signs) in that only the latter undergoes ‘cancellation’ in the construction of phrasal signs.

1.6.5 *Idiomatic Expressions*

Idioms also pose a potential problem for the strong locality claims made by SBCG, as I have outlined it here. It is well known that certain idiomatic interpretations arise only when the particular pieces of the idiom are in construction with one another. The proper characterization of the notion of ‘in construction with’, however, remains controversial. Since Nunberg et al. 1994, it has generally been agreed that syntactic flexibility is related to semantic decomposability. Thus a particularly decomposable idiom like *pull strings*, occurs flexibly in a variety of configurations, as illustrated in (44):

- (44) a. Sandy *pulled strings* to get Kim the job.
 b. *Strings* were *pulled* to get Kim the job.
 c. The *strings* that seem likely to have been *pulled* to get Kim the job were an offense to man and nature.
 d. We objected to the *strings* that Sandy had to *pull* to get Kim the job.
 e. Sandy *pulled* the *strings* that got Kim the job.
 f. The *strings* that Sandy *pulled*, nobody else could have *pulled*.

Idioms vary considerably in terms of their syntactic flexibility and it is perhaps unclear where to draw the line between an idiomatic sentence that should be allowed by the grammar and an extension of the grammar (or ‘language play’). However, it is reasonably clear that copredication is a necessary condition for idiomaticity. That is, in order for *pull strings* to receive its idiomatic interpretation, the second semantic argument of *pull* must also have *strings* predicated of it, however the grammar allows for that to happen.²⁵

²⁵ Sailer (2000) proposes a treatment of flexible idioms in terms of lexical constraints (called ‘conditions on lexical licensing’ (COLL)) that can access arbitrarily distant elements within a given phrasal structure. Sailer argues that the domain of COLL constraints should be the entire sentence (a sentential sign) in which the idiomatic word occurs. This is necessary, he claims, in order to describe what

My proposal, presented more fully in Sag to appear, uses the persistent defaults of Lascarides and Copestake (1999) to write lexical entries like those in (45): (LID is the feature LEXICAL-IDENTIFIER, originally proposed by Fillmore and Kay in unpublished work, and explained more fully in Sag 2007.)

$$(45) \left[\begin{array}{l} \text{FORM} \langle \textit{strings} \rangle \\ \text{SYN} \left[\begin{array}{l} \text{CAT} \left[\textit{noun} \right] \\ \text{LID} \boxed{\text{Q}}[\textit{strings_rel} /_p \textit{l_strings_rel}] \end{array} \right] \\ \text{VAL} \langle \rangle \\ \text{SEM} \left[\begin{array}{l} \text{INDEX } i \\ \text{RELS} \langle h_0:\boxed{\text{Q}}(i) \rangle \end{array} \right] \end{array} \right]$$

Assuming that literal and idiomatic relations are hierarchically organized as shown in Figure 1.20, then the noun *strings* will default to its literal interpretation except when its LID value is resolved to the idiomatic relation *i_strings_rel* by the idiomatic verb *pull*, whose lexical entry is sketched in (46):

$$(46) \left[\begin{array}{l} \text{FORM} \langle \textit{pull} \rangle \\ \text{SYN} \left[\text{VAL} \left\langle \left[\text{SYN NP}_i \right], \left[\text{LID } \textit{i_strings_rel} \right] \right\rangle \right] \\ \text{SEM} \left[\text{RELS} \langle h_0:\textit{i_pull_rel}(i, j) \rangle \right] \end{array} \right]$$

Making the reasonable assumption that the LID of a gap and its filler are identified in a filler-gap construction, it follows that the idiomatic resolution can take place in examples (44d-f), as well as (44a-c), thus solving what Nunberg et al. (1994) refer to as ‘McCawley’s Paradox’. This account of syntactically flexible, semantically decomposable idioms is fully compatible with the localist perspective of SBCG.

he takes to be purely syntactic restrictions on particular idiom ‘chunks’. I will not comment further on Sailer’s proposals here, or on the subsequent attempts to improve upon them by Soehn (2004, 2006). My approach differs from both of these in treating each idiom in terms of a single local constraint that interacts with other independently motivated aspects of the grammar.

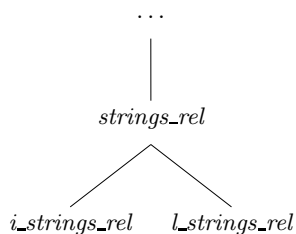


FIG. 1.20 Literal and Idiomatic Strings Relations

1.7 Conclusion

In this chapter, I have surveyed the issues of **locality of selection** and **locality of construction**, placing these matters in both theoretical and historical perspective. I have also sketched the basics of a particular blend of HPSG and Construction Grammar, called **Sign-Based Construction Grammar**, which draws a fundamental distinction between signs and constructs. The two basic principles of SBCG, the Subcategorization Principle and the Head Feature Principle are fundamental to the results presented here. These principles interact with appropriate lexical specifications and a particular inventory of features, including GAP, XARG, and LID, to make predictions of locality – a principled circumscription of the domains in which lexical selection and constructional constraints apply.

In addition, I have examined a number of problems involving nonlocal grammatical dependencies and have offered localist solutions to them. In the process, I have argued that SBCG offers numerous advantages, including:

- a unified approach to the apparent nonlocal dependencies involving embedded subjects and possessors (through the feature XARG),
- a principled account of the locality of lexical selection,
- a proper treatment of the syntactic flexibility of semantically decomposable idioms, including a resolution of ‘McCawley’s Paradox’,
- a simplification of the Subcategorization Principle (by eliminating the need for relational constraints), and
- a simplification of grammar rules (phrasal constructions) that precludes nonlocal constraints as a matter of principle.

Bibliography

- Allegranza, V. 1998. Determiners as functors: NP structure in Italian. In S. Balari and L. Dini (eds.), *Romance in Head-driven Phrase Structure Grammar*, pp 55–108, Stanford: CSLI Publications.
- Allegranza, V. 2007. *The Signs of Determination*. Frankfurt am Main: Peter Lang.
- Andrews, A. 1996. Semantic Case-Stacking and Inside-Out Unification. *Australian Journal of Linguistics* 16(1), 1–54.
- Asudeh, A. 2002. Richard III. In M. Andronis, E. Debenport, A. Pycha and K. Yoshimura (eds.), *Proceedings of the 38th Regional Meetings of the Chicago Linguistics Society*, Chicago: Chicago Linguistics Society.
- Bender, E. and Flickinger, D. P. 1999. Peripheral Constructions and Core Phenomena: Agreement in Tag Questions. In G. Webelhuth, J.-P. Koenig and A. Kathol (eds.), *Lexical and Constructional Aspects of Linguistic Explanation*, pp 199–214, Stanford: CSLI Publications.
- Bouma, G., Malouf, R. and Sag, I. A. 2001. Satisfying Constraints on Extraction and Adjunction. *Natural Language and Linguistic Theory* 1(19), 1–65.
- Chomsky, N. 1965. *Aspects of the Theory of Syntax*. MIT Press.
- Chomsky, N. 1970. Remarks on Nominalization. In R. Jacobs and P. S. Rosenbaum (eds.), *Readings in English Transformational Grammar*, Waltham, Massachusetts: Ginn and Co.
- Chung, S. 1982. Unbounded Dependencies in Chamorro Grammar. *Linguistic Inquiry* 13, 39–77.
- Chung, S. 1998. *The Design of Agreement: Evidence from Chamorro*. University of Chicago Press.
- Clements, G. N. 1984. Binding domains in Kikuyu. *Studies in the Linguistic Sciences* 14, 37–56.
- Copetake, A. 1992. The Representation of Lexical Semantic Information. Cognitive Science Research Papers 280, University of Sussex.
- Copetake, A., Flickinger, D., Pollard, C. J. and Sag, I. A. 2005. Minimal Recursion Semantics: an Introduction. *Research on Language and Computation* 3(4), 281–332.
- Copetake, A., Lascarides, A. and Flickinger, D. P. 2002. An Algebra for Semantic Construction in Constraint-Based Grammars. In *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics (ACL 2001), Toulouse, France*, University of Pennsylvania, Philadelphia: Association for Computational Linguistics.
- Daniels, M. and Meurers, D. 2004. GIDL: A Grammar Format for Linearization-Based HPSG. In S. Müller (ed.), *Proceedings of the HPSG-2004 Conference, Center for Computational Linguistics, Katholieke Universiteit Leuven*, pp 93–111, Stanford: CSLI

- Publications.
- Dench, A. and Evans, N. 1988. Multiple case-marking in Australian languages. *Australian Journal of Linguistics* 8(1), 1–47.
- Dickinson, M. 2004. Polish Numeral Phrases and Predicative Modification. To appear in *The Ohio State University Working Papers in Linguistics (OSUWPL)*.
- Diesing, M. 1990. Verb Movement and the Subject Position in Yiddish. *Natural Language and Linguistic Theory* 8, 41–79.
- Donohue, M. 2003. Review of Sandra Chung: The Design of Agreement: Evidence from Chamorro. *Linguistic Typology* 7, 285–292.
- Emonds, J. E. 1976. *A Transformational Approach to English Syntax: Root, Structure-Preserving, and Local Transformations*. New York: Academic Press.
- Flickinger, D. P. and Bender, E. M. 2003. Compositional Semantics in a Multilingual Grammar Resource. In E. M. Bender, D. P. Flickinger, F. Fouvry and M. Siegel (eds.), *Proceedings of the Workshop on Ideas and Strategies for Multilingual Grammar Development, ESSLLI 2003*, pp 33–42, Vienna, Austria.
- Gazdar, G. 1981. Unbounded Dependencies and Coordinate Structure. *Linguistic Inquiry* 12, 155–84.
- Gazdar, G. 1982. Phrase Structure Grammar. In P. Jacobson and G. K. Pullum (eds.), *The Nature of Syntactic Representation*, pp 131–186, Dordrecht: Reidel.
- Gazdar, G., Klein, E., Pullum, G. K. and Sag, I. A. 1985. *Generalized Phrase Structure Grammar*. Oxford: Basil Blackwell’s and Cambridge, MA: Harvard University Press.
- Gerds, D. and Hukari, T. 2001. A-Subjects and Control in Halkomelem. In D. Flickinger and A. Kathol (eds.), *The Proceedings of the 7th International Conference on Head-Driven Phrase Structure Grammar*, pp 100–123, Stanford: CSLI Publications.
- Harris, Z. S. 1946. From Morpheme to Utterance. *Language* 22(3), 161–183.
- Higgins, F. R. 1979. *The Pseudo-cleft Construction in English*. New York: Garland Publishers.
- Höhle, T. N. 1997. Vorangestellte Verben und Komplementierer sind eine Natürliche Klasse. In K. R. Durscheid, C. and M. Schwarz (eds.), *Sprache im Fokus: Festschrift für Heinz Vater zum 65. Geburtstag*, pp 107–120, Tübingen: Max Niemeyer Verlag.
- Hukari, T. E. and Levine, R. D. 1995. Adjunct extraction. *Journal of Linguistics* 31, 195–226.
- Joshi, A., Levy, L. and Takahashi, M. 1975. Tree Adjunct Grammars. *Journal Computer Systems Science* 10(1), 136–163.
- Kajita, M. 1968. *A Generative-Transformational Study of*

- Semi-Auxiliaries in Present-Day American English*. Tokyo: Sanseido.
- Kathol, A. 2000. *Linear Syntax*. New York, Oxford: Oxford University Press.
- Kay, P. 2002. English Subjectless Tagged Sentences. *Language* 78, 453–481.
- Kayne, R. and Pollock, J.-Y. 1978. Stylistic Inversion, Successive Cyclicity, and Move NP in French. *Linguistic Inquiry* 12, 93–133.
- Kiss, T. 1995. *Infinite Komplementation*. Tübingen: Max Niemeyer Verlag.
- Lascarides, A. and Copestake, A. 1999. Default Representation in Constraint-Based Frameworks. *Computational Linguistics* 25, 55–105.
- Levine, R. D. and Hukari, T. 2006. *The Unity of Unbounded Dependency Constructions*. Stanford University: CSLI Publications.
- Malouf, R. 2000. A head-driven account of long-distance case assignment. In R. Cann, C. Grover and P. Miller (eds.), *Grammatical Interfaces in HPSG*, pp 201–214, CSLI Publications.
- McCloskey, J. 1979. *Transformational syntax and model theoretic semantics: a case study in modern Irish*. Dordrecht, 1979: Reidel.
- McCloskey, J. 1990. Resumptive Pronouns, \bar{A} -Binding, and Levels of Representations in Irish. In R. Hendrick (ed.), *The Syntax of the Modern Celtic Languages*, pp 199–248, New York: Academic Press.
- Meurers, W. D. 1999. *Lexical Generalizations in the Syntax of German Non-Finite Constructions*. Ph. D.thesis, Seminar für Sprachwissenschaft, Universität Tübingen, Tübingen, Germany, published 2000 as Volume 145 in Arbeitspapiere des SFB 340, ISSN 0947-6954/00.
- Meurers, W. D. 2001. On Expressing Lexical Generalizations in HPSG. *Nordic Journal of Linguistics* 24(2).
- Müller, S. 2004. Continuous or Discontinuous Constituents? A Comparison between Syntactic Analyses for Constituent Order and Their Processing Systems. *Research on Language and Computation* 2(2), 209–257, Special Issue on Linguistic Theory and Grammar Implementation.
- Müller, S. 2005. Zur Analyse der deutschen Satzstruktur. *Linguistische Berichte* 201, 3–39.
- Nunberg, G., Sag, I. A. and Wasow, T. 1994. Idioms. *Language* 70(3), 491–538.
- Polinsky, M. 2007. Complementation in a Language Without Complement Clauses. Talk presented at the (Un)Usual Events Fest, Department of Linguistics, Stanford University.
- Pollard, C. J. 1994. Toward a Unified Account of Passive in German. In J. Nerbonne, K. Netter and C. J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar*, pp 273–296, Stanford

- University: CSLI Publications.
- Pollard, C. J. and Sag, I. A. 1987. *Information-based Syntax and Semantics, Vol. 1*. Stanford: CSLI Publications [Distributed by University of Chicago Press].
- Pollard, C. J. and Sag, I. A. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
- Potsdam, E. and Runner, J. T. 2001. Richard Returns: Copy Raising and its Implications. In *Proceedings of The Thirty-Seventh Regional Meeting of the Chicago Linguistics Society*, pp 453–468.
- Przepiórkowski, A. 1999. *Case Assignment and the Complement-Adjunct Dichotomy: A Non-Configurational Constraint-Based Approach*. PhD dissertation, Universität Tübingen.
- Pullum, G. K. and Zwicky, A. M. 1988. The syntax-phonology interface. In F. Newmeyer (ed.), *Linguistics: The Cambridge Survey; Volume I, Linguistic Theory: Foundations*, pp 255–280, Cambridge: CUP.
- Reape, M. 1994. Domain Union and Word Order Variation in German. In J. Nerbonne, K. Netter and C. J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar*, pp 151–197, Stanford University: CSLI Publications.
- Reape, M. 1996. Getting Things in Order. In H. Bunt and A. van Horck (eds.), *Discontinuous Constituency*, pp 209–253, Berlin, New York: Mouton de Gruyter.
- Rogers, A. 1974. A Transderivational Constraint on Richard? In *Papers from the Tenth Regional Meeting of the Chicago Linguistic Society*, pp 551–558, Chicago Linguistic Society.
- Sag, I. A. 1997. English Relative Clause Constructions. *Journal of Linguistics* 33(2), 431–484.
- Sag, I. A. 2001. Dimensions of Natural Language Locality. Unpublished paper presented at the Eighth Annual Conference on Head-Driven Phrase Structure Grammar. NTNU. Trondheim, Norway.
- Sag, I. A. 2007. Sign-Based Construction Grammar: An informal synopsis. Unpublished ms, Stanford University.
- Sag, I. A. to appear. Idioms by Persistent Default, unpublished manuscript, Stanford University. Based on a talk presented at Collocations and Idioms 2006: Linguistic, Computational, and Psycholinguistic Perspectives, Berlin, Germany.
- Sag, I. A. and Pollard, C. J. 1991. An Integrated Theory of Complement Control. *Language* 67, 63–113.
- Sag, I. A., Wasow, T. and Bender, E. M. 2003. *Syntactic Theory: A Formal Introduction*. Stanford: CSLI Publications, second edition.
- Sailer, M. 2000. *Combinatorial Semantics and Idiomatic Expressions in Head-Driven Phrase Structure Grammar*. Ph.D.thesis, Eberhard-Karls-Universität Tübingen.

- Soehn, J.-P. 2004. License to COLL. In S. Müller (ed.), *Proceedings of the HPSG-2004 Conference, Center for Computational Linguistics, Katholieke Universiteit Leuven*, pp 261–273, Stanford: CSLI Publications.
- Soehn, J.-P. 2006. On Idiom Parts and their Contexts. *Linguistik online* 27(2), 11–28.
- Steedman, M. 1996. *Surface Structure and Interpretation*. Cambridge, MA: MIT Press.
- Steedman, M. 2000. *The Syntactic Process*. Cambridge, MA: MIT Press/Bradford Books.
- Thráinsson, H. 1979. *On Complementation in Icelandic*. New York: Garland Publishing.
- Torrego, E. 1984. On Inversion in Spanish and Some of Its Effects. *Linguistic Inquiry* 15, 103–129.
- Van Eynde, F. 1998. The Immediate Dominance Schemata of HPSG. A deconstruction and a reconstruction. In P. Coppen, H. van Halteren and L. Teunissen (eds.), *Computational Linguistics in the Netherlands 1997. Selected Papers from the Eighth CLIN Meeting*, pp 119–133, Amsterdam/Atlanta: Rodopi.
- Van Eynde, F. 2006. NP-Internal Agreement and the Structure of the Noun Phrase. *Journal of Linguistics* 42(1), 139–186.
- Van Eynde, F. 2007. The Big Mess Construction. In S. Müller (ed.), *The Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar, Stanford University*, pp 415–433, Stanford: CSLI Publications.
- Zaenen, A. 1983. On syntactic binding. *Linguistic Inquiry* 14, 469–504.
- Zec, D. 1987. On Obligatory Control in Clausal Complements. In M. Iida, S. Wechsler and D. Zec (eds.), *Working Papers in Grammatical Theory and Discourse Structure*, pp 139–168, Stanford: CSLI Publications.