Remarks on locality

Ivan A. Sag
Stanford University

Proceedings of the 14th International Conference on
Head-Driven Phrase Structure Grammar

Stanford Department of Linguistics and CSLI's LinGO Lab

Stefan Müller (Editor)

2007

Stanford, CA: CSLI Publications

pages 394–414

Sag, Ivan A. 2007. Remarks on locality. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar, Stanford Department of Linguistics and CSLI's LinGO Lab*, 394–414. Stanford, CA: CSLI Publications. DOI: 10.21248/hpsg.2007.23.

Abstract

This paper proposes a modification of HPSG theory – Sign-Based Construction Grammar – that incorporates a strong theory of both selectional and constructional locality. A number of empirical phenomena that give the appearance of requiring nonlocal constraints are given a principled, localist analysis consistent with this general approach, which incorporates certain insights from work in the tradition of Berkeley Construction Grammar, as exemplified by Fillmore et al. (1988), Kay and Fillmore (1999), and related work.

1 Introduction

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.[1] For example, while it is commonplace to find a language containing a verb like *go*, which allows a directional PP complement, but not a NP object, there are no languages (as far as we know) where we find a verb like *go* that imposed the same requirement on the complementation pattern realized within its sentential complement. That is we would not expect to find a verb *og* whose selectional properties produced contrasts like the following:

(1) a. Lee **oged** that someone ran into the room.

b. *Lee **oged** that someone proved a theorem.

The question of locality of subcategorization seems to have fallen by the wayside within mainstream generative grammar. It is important to realize, however, that '$\overline{\text{X}}$ Theory', as first developed in Chomsky 1970 (but cf. Harris 1946), bears on this question. A verb that selects an NP complement (a transitive verb) is really selecting for a phrase with a (nonpredicative) nominal head. And $\overline{\text{X}}$ 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

[†]Some of the ideas developed here were first presented at the 2001 HPSG Conference, held at NTNU in Trondheim, Norway. I would like to thank Emily Bender, 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 for valuable discussions about locality. I am particularly grateful to Doug Ball, Detmar Meurers and Stefan Müller for detailed comments on an earlier draft of this paper. 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]The locality of selection is one of the theoretical issues that were hotly debated during the 1960s. For further discussion and historical review, see Sag to appear a.

projection (e.g. the maximal NP headed by a given noun, the maximal AP headed by a given adjective, etc.). $\overline{X}$ Theory thus plays a crucial role in considerations of locality – a verb refers to the category features of the phrases it combines with, i.e. the phrases (NP, AP, etc.) that are sisters of the verb and it follows that those phrases will be headed by a word of the appropriate syntactic category.

These ramifications of $\overline{X}$ Theory played an important role in the development of Generalized Phrase Structure Grammar (GPSG). Gazdar (1981) and Gazdar et al. (1985) argued that $\overline{X}$ 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 such features as CASE, VFORM, NFORM, PFORM, PRED, AUX, and SLASH. With this feature inventory, the explanatory domain of $\overline{X}$ 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 the values for these features are 'percolated up' from lexical heads to the phrases they project (by the Head Feature Principle (HFP), an uncontroversial principle of $\overline{X}$ Theory), the information relevant to all these phenomena becomes locally accessible to the lexical items that combine with those phrasal projections.

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 (2):

(2) 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 HPSG) 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.

Closely related to selectional locality is the issue of **locality of construction** – the problem of delimiting the syntactic and semantic information accessible to

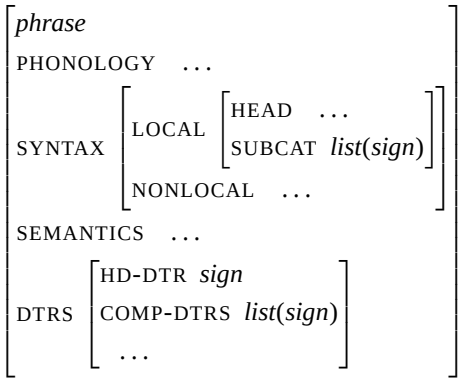


Figure 1: Feature Geometry of Pollard and Sag 1987

grammar rules. That is, just as we observe empirically that there are no languages with extended subcategorization of the sort illustrated in (1) above, I would argue that there are also no languages where one must propose a grammar rule that directly relates two elements across clauses. In all apparent cases of this that I am familiar with, there is a satisfying feature-based analysis of the construction in question that conforms to a strict localist architecture.

2 The SYNSEM Locality Hypothesis

The feature geometry proposed by Pollard and Sag (1987) [henceforth P&S-87] (sketched in Figure 1, taken together with their Subcategorization Principle in (3)), failed to place sufficient constraints on which elements could be selected by a given word.[2]

(3) **Subcategorization Principle** (P&S-87: 71):

$$\left[\text{DTRS } \textit{head-struc}\right] \Rightarrow \left[\begin{array}{l}\text{SYN|LOC|SUBCAT } \boxed{A} \\ \text{DTRS } \left[\begin{array}{l}\text{HD-DTR } \left[\text{SYN|LOC|SUBCAT } \boxed{A} \oplus \boxed{B}\right] \\ \text{COMP-DTRS } \boxed{B}\end{array}\right]\end{array}\right]$$

In this set-up, since phrasal signs have daughters, the elements on a verb's SUBCAT list do too. Hence a lexical entry could easily be written for a verb that is subcategorized for a VP complement that must contain a direct object NP or (even more permissively) for an S whose VP contained an S whose VP contained an object specified as, say, [CASE dative]. Early HPSG thus embodied little in the way of a theory of subcategorization locality.

[2]For uniformity of presentation, I here reverse the order of elements on SUBCAT lists from that assumed in P&S-87. The symbol '$\oplus$' denotes list concatenation (also referred to as the 'addition' or the 'appending' of two lists.)

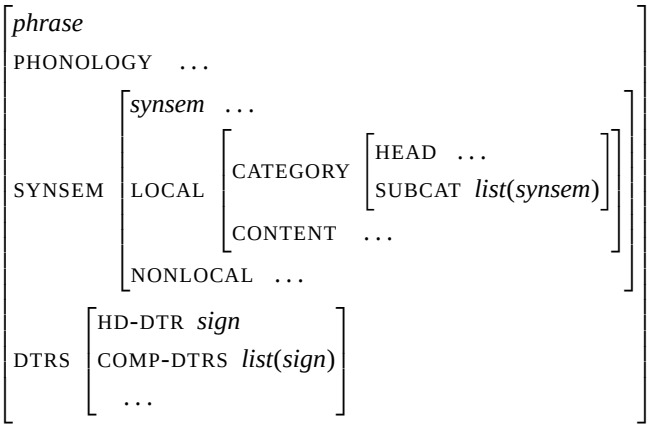


Figure 2: Feature Geometry of Pollard and Sag 1994

The proposals made by Pollard and Sag (1994) [henceforth P&S-94] embodied an attempt to remedy this defect. By introducing the feature SYNSEM and the syntactico-semantic complexes ('synsem objects') that served as values of SYNSEM, P&S-94 were able to limit the information that was accessible under lexical selection, as shown in Figure 2.

This feature geometry worked together with a revised Subcategorization Principle, formulated in (4):[3]

(4) **Subcategorization Principle** (a formalization of P&S-94: 34):

$$\left[\text{DTRS } \textit{head-struc}\right] \Rightarrow \left[\begin{array}{l}\text{SS|LOC|CAT|SUBCAT } \boxed{A} \\ \text{DTRS } \left[\begin{array}{l}\text{HD-DTR|SS|LOC|CAT|SUBCAT } \boxed{A} \oplus \texttt{s2s}(\boxed{B}) \\ \text{COMP-DTRS } \boxed{B}\end{array}\right]\end{array}\right]$$

We may refer to the feature geometry in Figure 2, taken together with the Subcategorization Principle in (4), as the SYNSEM Locality Hypothesis (SSLH).[4]

The SSLH ensures that if a lexical entry includes a constraint on a member of the SUBCAT list, that constraint will apply to the SYNSEM value of the corresponding valent (subject, complement, or specifier) that that word cooccurs with. There is no direct access to information about any element that appears within those valents, e.g. a direct object within a VP complement, or an object within a sentential complement of a sentential complement. There is only indirect access to such elements whenever certain SYNSEM properties of a given valent are determined by or correlated with those of some element it contains.

[3]The function `s2s` (`signs-to-synsems`) maps a list of signs onto the corresponding list of synsem objects.

[4]The SSLH also includes the prediction that (morpho-)phonological information is unavailable for lexical selection. Space limitations prevent a proper evaluation of this independent issue here.

The SSLH embodies a quite particular claim: taken together with a theory of what SYNSEM values are, it ensures that the grammatical constraints that concern the following phenomena all function within the same locality domain: **category selection** (strict subcategorization in Chomsky's sense), **case assignment**, **government** (of the form of a complement's head), and **(non-anaphoric) agreement**. In many clear cases, these predictions are correct, though there remain certain issues of controversy, some of which I discuss below.

Note that under these assumptions it is not possible to write a lexical entry that selects for 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 always determined by general constraints on the 'inheritance' of SLASH specifications.[5]

I note in passing that the hypothesis that information about filler-gap dependencies should be locally encoded has been confirmed now by evidence from numerous languages. All of the following phenomena, for example, are sensitive to the presence of a filler-gap dependency and are easily described given the localist, feature-based approach to unbounded dependencies pioneered in GPSG/HPSG research: Irish Complementizers, 'Stylistic' Inversion (Romance), Kikuyu Downstep Suppression, Austronesian Verb Agreement, Yiddish Inversion, Icelandic Expletives, Thompson Salish Verb Morphology, Adyghe '*wh*-agreement'.[6]

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) declared appropriate for the type *phrase*. Grammar schemata were introduced in PS-94 as the HPSG analog of grammar rules. These schemata specified an inventory of phrase types, where phrases had the geometry shown in Figure 2 above. Since phrases contained daughter structures of arbitrary depth and schemata imposed constraints directly on phrases, there was nothing in this set-up that imposed any notion of locality. 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:

[5] 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'), one could write a lexical entry that forced that gap to appear at some fixed depth within the infinitival complement of *hard*, an expressible, but cross-linguistically non-occurring possibility.

[6] For further discussion, see Hukari and Levine 1995, Levine and Hukari 2006, Sag to appear a, and the references cited there.

> [...] 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] 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:

(5) 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 (6), which argues that the agreement in question is semantic, rather than syntactic:[9]

(6) a. Sears is open, aren't they?

b. At least one of us is sure to win, aren't we?

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 two elements that are not sisters, as shown in Figure 3.

As Bender and Flickinger argue, the English tag-question construction argues 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, i.e. to the S. 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 values of the subject-encoding feature of the main clause and that of the tag clause (the clauses that are shaded in Figure 3).

[7] See, for example, Reape (1994, 1996), Kathol 2000, and Daniels and Meurers 2004.

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

[9] See Kay 2002 and the references cited there.

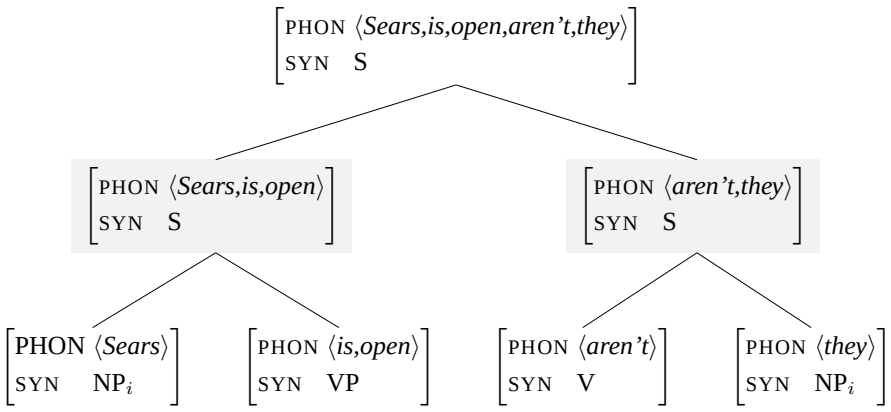


Figure 3: A Tag-Question

4 Signs, Constructions, and Constructs

I propose to modify HPSG theory so as to incorporate the strong constraints of the actual practice of the HPSG community. To this end, phrases should not be endowed with the feature DTRS. Phrases, like words, specify values for the features PHONOLOGY, SYNTAX, and SEMANTICS. Second, signs should be distinguished from the constructions that license them. (What I mean by this will become clear in a moment.)

A construction, like a schema in PS-94, is intuitively a constraint defining 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 and in addition puts constraints on the sign that results from such a combination. On this conception, 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'.

Notice that we may now return to a simpler feature geometry like the one in PS-87, eliminating the feature SYNSEM. In addition, with no distortion of the grammar's intended effect, we may reformulate constructs as feature structures, as shown in (7):[10] This last move is in fact easily achieved by the type declarations sketched in Figure 5, which define part of the type hierarchy shown in Figure 6:

Of course, this system of grammar doesn't define complex expressions until

[10]For expositional purposes, I will sometimes represent constructs in tree notation and will use SYNTAX and SEMANTICS values, as in Figure 4.

$$\begin{bmatrix} \textit{phr-cxt} & \\ \text{MTR} & \begin{bmatrix} \textit{phrase} & \\ \text{PHON} & \langle\, \textit{Kim}\ ,\ \textit{walks}\, \rangle \\ \text{SYN} & \text{S} \\ \text{SEM} & \textbf{walk}\textit{(}\textbf{k}\textit{)} \end{bmatrix} \\ \text{DTRS} & \left\langle \begin{bmatrix} \text{PHON} & \langle\, \textit{Kim}\, \rangle \\ \text{SYN} & \text{NP} \\ \text{SEM} & \textbf{k} \end{bmatrix}, \begin{bmatrix} \text{PHON} & \langle\, \textit{walks}\, \rangle \\ \text{SYN} & \text{V} \\ \text{SEM} & \textbf{walk} \end{bmatrix} \right\rangle \end{bmatrix}$$

Figure 4: A Clausal Construct

$$\textit{cxt}: \begin{bmatrix} \text{MOTHER} & \textit{sign} \\ \text{DTRS} & \textit{list(sign)} \end{bmatrix}$$

$$\textit{ph-cxt}: \begin{bmatrix} \text{MOTHER} & \textit{phrase} \end{bmatrix}$$

$$\textit{hd-cxt}: \begin{bmatrix} \text{HD-DTR} & \textit{sign} \end{bmatrix}$$

$$\textit{sign}: \begin{bmatrix} \text{PHON} & \textit{list(phon-structure)} \\ \text{FORM} & \textit{list(morph-form)} \\ \text{SYNTAX} & \textit{syn-obj} \\ \text{SEMANTICS} & \textit{sem-obj} \end{bmatrix}$$

Figure 5: Type Declarations

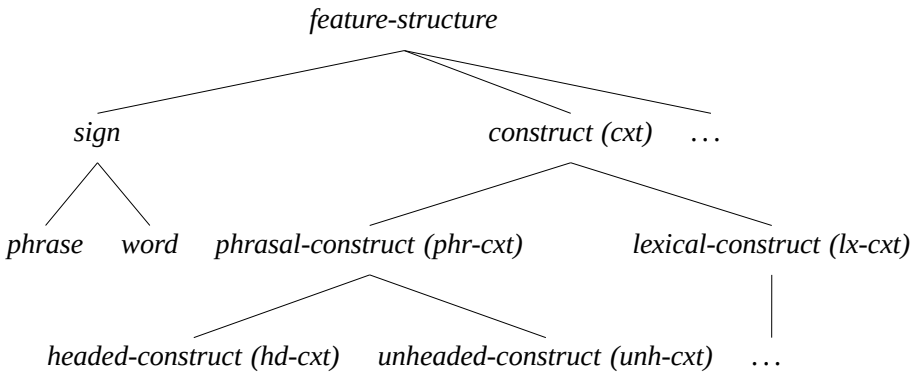


Figure 6: A SBCG Type Hierarchy

we include a principle like the the following, which allows recursive application of constructions:

(7) **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.

I will refer to any framework that draws the distinction between signs and constructs as **Sign-Based Construction Grammar** (**SBCG**),[11] though of course this is still a kind of HPSG, given that it embodies signs, linguistically motivated types, type constraints, and a hierarchically organized lexicon, *inter alia*.

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 linguistic descriptions. It also fosters the development of general principles constraining the distribution of feature specifications across constructs. In fact, the fundamental principles of P&S-94 are now recast as constraints on constructions, as shown in (8):[12]

(8) **Head Feature Principle:**

$$\textit{hd-cxt} \Rightarrow \begin{bmatrix} \text{MTR} & \left[\text{SYN}|\text{CAT} \ \boxed{1}\right] \\ \text{HD-DTR} & \left[\text{SYN}|\text{CAT} \ \boxed{1}\right] \end{bmatrix}$$

Subcategorization Principle:

$$\textit{hd-cxt} \Rightarrow \begin{bmatrix} \text{MTR} & \left[\text{SYN}|\text{VAL} \ \boxed{A}\right] \\ \text{DTRS} & \boxed{B} \bigcirc \langle \boxed{1} \rangle \\ \text{HD-DTR} & \boxed{1}\left[\text{SYN}|\text{VAL} \ \boxed{A} \oplus \boxed{B}\right] \end{bmatrix}$$

Note that the Subcategorization Principle is stated here without appeal to the *signs-to-synsems* relation.

Finally, this proposal also provides a new way of making sense of lexical rules, i.e. by treating them as varieties of lexical construction. We may posit three sub-

[11] For an early formulation, see Chapter 16 of Sag, Wasow, and Bender 2003. Here I follow the detailed presentation of SBCG in Sag 2007, where various features (e.g. SYNSEM, LOCAL, NONLOCAL, HEAD) are eliminated and SUBCAT is replaced by VALENCE (VAL).

[12] '$\bigcirc$' is Reape's domain union operator: '$\boxed{A} \bigcirc \boxed{B}$' is satisfied by any list containing exactly the elements of $\boxed{A}$ and $\boxed{B}$, as long as any α which precedes some β in $\boxed{A}$ or in $\boxed{B}$ also precedes β in $\boxed{A} \bigcirc \boxed{B}$. '$\bigcirc$' is thus a 'shuffle' operator.

types 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 Chap 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 type *lexeme*; inflectional constructions involve unary constructs (of type *infl-cxt*) whose mother is of type *word* and whose daughter is of type *lexeme*; and post-inflectional constructions involve unary constructs (of type *post-infl-cxt*) 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.

5 Some Analytic Issues

The SBCG framework is attractive for its simplicity and strong predictive power. However, its predictions may be too strong, as there remain various empirical phenomena that, at least in their outward appearance, appear to defy the localism embodied in SBCG. In the remainder of this paper, I will examine a number of such phenomena, showing that an attractive localist analysis is available.

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

(9) a. I prefer [for [*they to be happy]]

b. I prefer [for [them to be happy]].

But given the standardly assumed structure in (9), 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 (10), that subject must be compatible with nominative case:

(10) [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 7.[13] Assuming this structure, rather than the one in (9), 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.

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

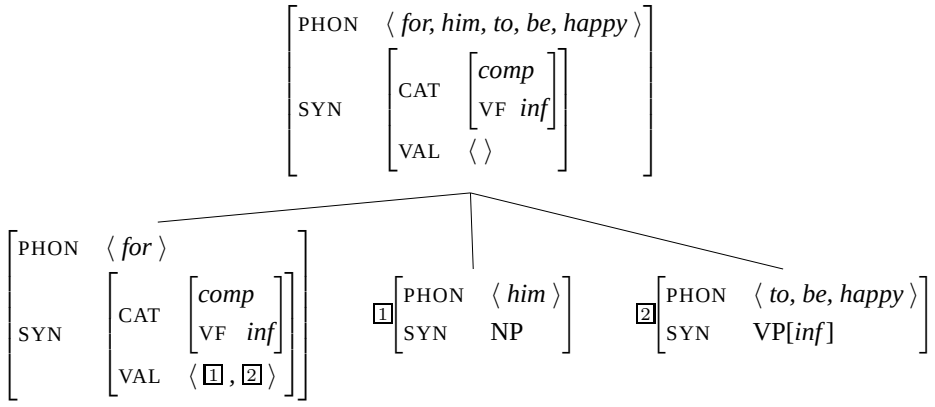


Figure 7: A *For-To* Clause

Moreover, the structure in (10) is independently motivated, for it provides an immediate account of contrasts like the following, first noted by Emonds (1976):

(11) 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 (11a-b), there is no constituent for the italicized modifiers to modify in (11c-d). The deviance of these examples follows from the same constraints that disallow the indicated modifiers in (12a-b), whose structure is analogous to the new *for/to*-clausal structure:

(12) a. *Kim persuaded *in St. Louis* Sandy to rent a house cheap.

b. *Lee believed *in these circumstances* Sandy to be in the right.

5.2 Case Stacking Languages

One of the best-known examples of apparent nonlocal case assignment come from languages that allow case 'stacking', as in the following examples from Martuthunira, a Pama-Nyungan language:

(13) 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))

(14) 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 (13)–(14) seems to pose a serious challenge for any locality hypothesis, and certainly for the SSLH.

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'[14] looks like (15):

(15)
$$\begin{bmatrix} \textsc{phon} & \langle\, \textit{tharnt-} \,\rangle \\ \textsc{syn} & \begin{bmatrix} \textsc{cat} & \begin{bmatrix} \textit{noun} \\ \textsc{case} & \boxed{B} \end{bmatrix} \\ \textsc{val} & \left\langle \textsc{np}[\textsc{case}\ \langle \textit{prop} \rangle \oplus \boxed{B}\,] \right\rangle \end{bmatrix} \end{bmatrix}$$

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 (16):

(16)
$$\begin{bmatrix} \textsc{phon} & \langle\, \textit{nhuwalalha} \,\rangle \\ \textsc{syn} & \begin{bmatrix} \textsc{cat} & \begin{bmatrix} \textit{verb} \\ \textsc{case} & \boxed{B}\langle\,\rangle \end{bmatrix} \\ \textsc{val} & \left\langle \textsc{np}[\langle \textit{nom} \rangle \oplus \boxed{B}],\ \textsc{np}[\langle \textit{acc} \rangle \oplus \boxed{B}] \right\rangle \end{bmatrix} \end{bmatrix}$$

Finite verbs bear an empty CASE specification. However, (16) 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 8.

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

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

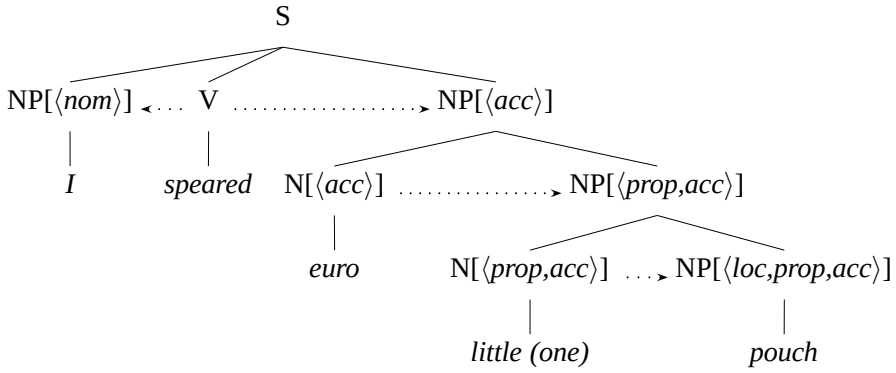


Figure 8: Case Government in Martuthunira

given head-dependent combination. The morphological case inflections are based on local CASE specfications, just as they are in languages that lack case stacking. But when multiple case affixes are present (e.g. on *pouch* in (16), 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 (16) would fail to meet the VALENCE requirements of the verb *speared*. 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.

5.3 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 (17):

(17) 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 (Gerdts 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 (18):

(18) a. NP_i promise [CMP he_i VP]

b. NP persuade NP$_i$ [CMP 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 (19):

(19) 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.[15]

However, it would be desirable to use the same feature to make genitive pronouns that are 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).[16] Because XARG is a category feature, it percolates information about a designated phrasal constituent, as illustrated in Figure 9.

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 (17) above can be treated simply by associating the relevant lexical entry for *looks* (*like*) with the VAL list in (20):

(20) $\left[\text{VAL} \left\langle \text{NP}_i\ , \left[\begin{matrix}\text{S} \\ \text{XARG}\ \ \text{NP}[\textit{pro}]_i\end{matrix}\right] \right\rangle\right]$

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

[16] Sag and Pollard (1991) proposed a semantic feature EXTERNAL-ARGUMENT (XARG), which makes 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).

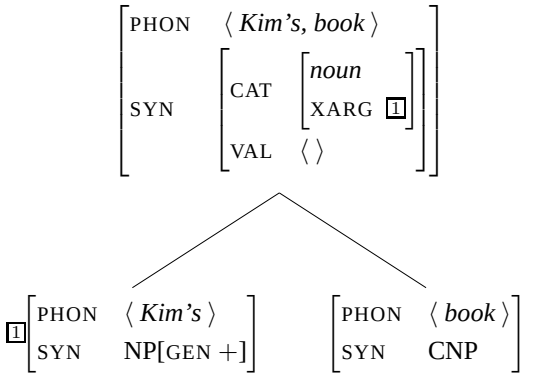


Figure 9: XARG Analysis of Genitive-Embedding NP

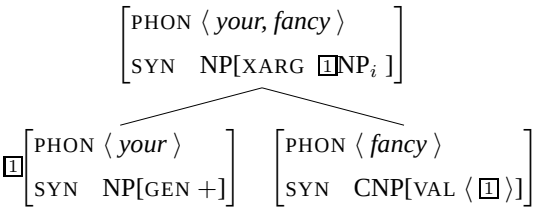
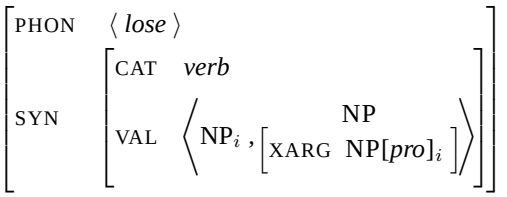


Figure 10: XARG Analysis of *your fancy*

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 (21):

(21)


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 10.

All of the phenomena just enumerated, in addition to the 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.

5.4 Idiomatic Expressions

Idioms also potentially pose a locality issue. 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 (22):

(22) 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.[17]

My proposal, presented more fully in Sag to appear b, uses the persistent defaults of Lascarides and Copestake (1999) to write lexical entries like those in (23) (LID is the feature LEXICAL-IDENTIFIER explained more fully in Sag 2007):

[17] 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 he takes to be purely syntactic restrictions on particular idiom 'chunks'. Space limitations prevent me from providing a fuller discussion of Sailer's proposals, or the subsequent attempts to improve upon them by Soehn (2004, 2006). My approach differs from these in treating each idiom in terms of a single local constraint that interacts with other aspects of the grammar.

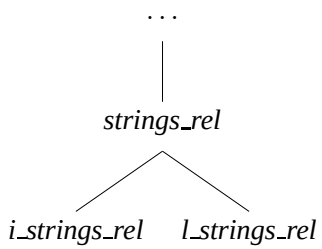


Figure 11: Literal and Idiomatic Strings Relations

(23)
$$\begin{bmatrix} \text{PHON} & \langle\ \textit{strings}\ \rangle \\ \text{SYN} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \textit{noun} \\ \text{LID} & \boxed{0}[strings\_rel\ /_p\ l\_strings\_rel] \end{bmatrix} \\ \text{VAL} & \langle\ \rangle \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{INDEX}\ i \\ \text{RELS}\ \langle\ h_0{:}\boxed{0}(i)\ \rangle \end{bmatrix} \end{bmatrix}$$

Assuming that literal and idiomatic relations are hierarchically organized as shown in Figure 11, 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 lexical entry for the idiomatic verb *pull*, whose lexical entry is sketched in (24):

(24)
$$\begin{bmatrix} \text{PHON} & \langle\ \textit{pull}\ \rangle \\ \text{SYN} & \begin{bmatrix} \text{VAL} & \left\langle \begin{bmatrix} \text{SYN}\ \text{NP}_i \end{bmatrix}, \begin{bmatrix} \text{LID}\ i\_strings\_rel \\ \text{SYN}\ \text{NP}_j \end{bmatrix} \right\rangle \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{RELS}\ \langle\ h_0{:}i\_pull\_rel(i,j)\ \rangle \end{bmatrix} \end{bmatrix}$$

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 (22d-f), as well as (22a-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.

6 Conclusion

In this paper, I have surveyed and offered localist solutions to a number of problems involving nonlocal grammatical dependencies. I have proposed a version of HPSG theory – Sign-Based Construction Grammar – that is based on a distinction

between signs and constructs. Drawing the distinctions in the way I have outlined provides numerous advantages, including the following:

- Solutions are offered to a number of problems not solved by previous versions of HPSG (e.g. Pollard and Sag 1994 or Ginzburg and Sag 2000).
- Lexical selection is localized in a principled fashion.
- Previous results in HPSG are preserved.
- Principles, e.g. the Subcategorization Principle, are simplified, e.g. by eliminating the need for relational constraints such as `signs-to-synsems`.
- Phrasal schemata (constructions) are localized, i.e. they are fundamentally like CFG grammar rules.

References

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.

Chomsky, N. 1970. Remarks on Nominalization. In R. Jacobs and P. S. Rosenbaum (eds.), *Readings in English Transformational Grammar*, Waltham, Massachusetts: Ginn and Co.

Copestake, A. 1992. The Representation of Lexical Semantic Information. Cognitive Science Research Papers 280, University of Sussex.

Daniels, M. and Meurers, D. 2004. GIDLP: 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.

Dickinson, M. 2004. Polish Numeral Phrases and Predicative Modification. To appear in *The Ohio State University Working Papers in Linguistics (OSUWPL)*.

Emonds, J. E. 1976. *A Transformational Approach to English Syntax: Root, Structure-Preserving, and Local Transformations*. New York: Academic Press.

Fillmore, C., Kay, P. and O'Connor, M. 1988. Regularity and Idiomaticity in Grammatical Constructions: The Case of *Let Alone*. *Language* 64, 501–538.

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., Klein, E., Pullum, G. K. and Sag, I. A. 1985. *Generalized Phrase Structure Grammar*. Oxford: Basil Blackwell's and Cambridge, MA: Harvard University Press.

Gerdts, 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.

Ginzburg, J. and Sag, I. A. 2000. *Interrogative Investigations: the form, meaning, and use of English Interrogatives*. Stanford, California: CSLI Publications.

Harris, Z. S. 1946. From Morpheme to Utterance. *Language* 22(3), 161–183.

Höhle, T. N. 1997. Vorangestellte Verben und Komplementierer sind eine Naturliche Klasse. In K. R. Durscheid, C. and M. Schwarz (eds.), *Sprache im Fokus: Festschrift für Heinz Vater zum 65. Geburtstag*, pp 107–120, Max Niemeyer Verlag.

Hukari, T. E. and Levine, R. D. 1995. Adjunct extraction. *Journal of Linguistics* 31, 195–226.

Kay, P. 2002. English Subjectless Tagged Sentences. *Language* 78, 453–481.

Kay, P. and Fillmore, C. 1999. Grammatical Constructions and Linguistic Generalizations: The *What's X Doing Y?* Construction. *Language* 75, 1–33.

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.

Meurers, W. D. 1999. German Partial-VP Fronting Revisited. In G. Webelhuth, J.-P. Koenig and A. Kathol (eds.), *Lexical and Constructional Aspects of Linguistic Explanation*, Stanford: CSLI Publications, (revised version of the paper presented at the 3rd Int. Conference on HPSG, Marseille, May 1996).

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.

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.

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. 2007. Sign-Based Construction Grammar: An informal synopsis. Unpublished ms, Stanford University.

Sag, I. A. to appear a. Feature Geometry and Predictions of Locality. In G. Corbett and A. Kibort (eds.), *Proceedings of the Workshop on Features*, Oxford: Oxford University Press.

Sag, I. A. to appear b. 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-Universitt 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.

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 Structu re*, pp 139–168, Stanford: CSLI Publications.