14

Coordination and Underspecification

Ivan A. Sag

14.1 Introduction

Coordinate structures have posed a serious problem for HPSG ever since the idea that models of linguistic objects are 'complete' (i.e. totally well-typed and sort-resolved) became a standard assumption more than a decade ago. The problem, *tout court*, is the question of what feature structure to associate with the mother of the bracketed coordinate structure in examples like the following, where the categories of the conjuncts differ as indicated:

(1) Pat is [wealthy and a Republican]. [AP & NP] (Sag et al. 1985)

(2) Kim [likes bagels and is happy]. [[AUX −] & [AUX +]]

(3) Er [findet und hilft] Frauen.
OBJ.ACC OBJ.DAT
He finds and helps women
"He finds and helps women."
[[COMPS ⟨[*acc*]⟩] & [COMPS ⟨[*dat*]⟩]] (Ingria 1990)

(4) I certainly will, and you already have, {*clarify/*clarified the situation}
{set the record straight} with respect to the budget.
[[SLASH? {VP[*bse*]}] & [SLASH? {VP[*psp*]}]] (Pullum and Zwicky 1986)

The Proceedings of the 9th International Conference on HPSG.
Jong-Bok Kim and Stephen Wechsler.

(5) Kogo/*Co [Janek lubi a Jerzy
(acc/gen)/(nom/acc) obj.acc
who John likes and George
nienawidzi]?
obj.gen
hates

"Who/*What does John like and George hate?"
[[slash {NP[*acc*]}] & [slash {NP[*gen*]}]] (Polish: Dyła 1984)

(6) Dajcie [wina i całą świnię]!
give wine.gen and whole.acc pig.acc

"Serve some wine and a whole pig!"
[[case *gen*] & [case *acc*]] (Polish: Przepiórkowski 1999)

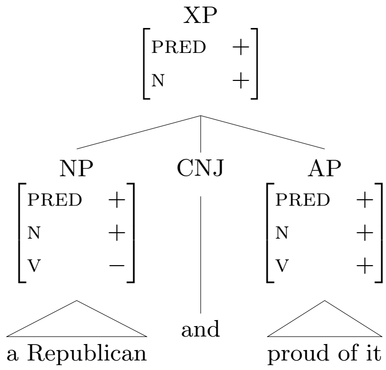
Various empirical problems that bear on this issue have been noted by Zaenen and Karttunen (1984), Pullum and Zwicky (1986), Jacobson (1987), and Ingria (1990), among others. In this paper I explore the idea that most (if not all) of these problems can be dealt with in HPSG by a simple change to the framework's foundational assumptions. I will consider an approach to these problems that involves suspending the requirement that feature structures be 'sort-resolved'.

14.2 Background

Work in Generalized Phrase Structure Grammar (GPSG),[1] included the proposal that a coordinate mother's feature structure is determined from the feature structures of its conjuncts via a set-theoretic relation (intersection of sets of feature-value specifications), an idea later adapted (in terms of union of atomic values) by Dalrymple and Kaplan (2000). The GPSG analysis treated (1) in terms of partially specified feature structures like the one labelling the mother node in (7):

[1] Sag et al. 1985; Gazdar et al. 1985.

(7)



Obviously, this proposal relied crucially on the assumption that well-formed feature structures need not be fully resolved.

As observed by Jacobson (1987), however, the GPSG analysis encounters difficulty in dealing with contrasts like the following:

(8) a. Kim grew wealthy.
b. *Kim grew a Republican.
c. Kim grew and remained wealthy.
d. *Kim grew and remained a Republican.
e. *Kim grew and remained [wealthy and a Republican].

That is, assuming that *grow* selects for an [N +, V +] (AP) complement and that *remain* is freer, requiring only that its complement be [V +] (AP or NP), the GPSG treatment predicts that the coordinate verb [*grew and remained*] should, like *remain*, impose only the weaker requirement on the complement. But this cannot be right, Jacobson argues, because [V +] is the category that would be associated with phrases like *wealthy and a Republican* (cf. (1)) and such phrases cannot serve as the complement of *grew and remained*, as (8e) shows.

Bayer and Johnson (1995) and Bayer (1996) propose a solution to this and related problems in terms of Type Logical Grammar, a species of Categorial Grammar where functional categories correspond to implication. On such an approach, an expression of type `VP/NP` is one that can give rise to a `VP` IF it is 'given' an `NP`. Implicational categories and conjunctive/disjunctive categories (categories which are built up with ∧ and ∨ and which can be simplified according to familiar logical principles, e.g. '∧-Elimination') interact in such a way as to solve Jacobson's puzzle. That is, if *remained* is of type `VP/(NP∨AP)`, then the logic of categories allows us to infer that *remained* may also be of type `VP/AP`, which allows it to coordinate with *grew*, which is of that type (and only of that type). The resulting coordinate verb, assuming that coordination requires category identity, is also VP/AP. This correctly accounts for (8a-d). In order for expressions of unlike category to coordinate, they must each be weakened to a 'lowest common denominator'. Thus, we may infer that an expression of type NP or one of type AP also leads a life as an NP$\vee$AP expression and this is the only type that can be assigned to a coordinate expression like *wealthy and a Republican*. But this cannot combine with [*grew and remained*], whose type is VP/AP, as we just saw. Thus (8e) is correctly ruled out – for the same reason that r cannot be derived from the premises $p \vee q$ and $p \rightarrow r$.

Similarly, Bayer (and Johnson) provide a solution for some of the other examples noted in section 1. For example, by analyzing a case-syncretic noun as an expression of a case-conjunctive type (that in many circumstances gives rise to expressions of a simpler type via $\wedge$-Elimination), one arrives at an analysis of the coordination of verbs selecting objects with distinct case, as in the following derivation of (3) above (after Bayer 1996):

(9)

findet	und	hilft	Frauen
VP/NP$\boxed{\text{A}}$	$(\alpha/_{\text{L}}\alpha)/\alpha$	VP/NP$\boxed{\text{D}}$	NP$\boxed{\text{N}}\wedge$NP$\boxed{\text{A}}\wedge$NP$\boxed{\text{D}}\wedge$NP$\boxed{\text{G}}$
VP/(NP$\boxed{\text{A}}\wedge$NP$\boxed{\text{D}}$)		VP/(NP$\boxed{\text{A}}\wedge$NP$\boxed{\text{D}}$)	NP$\boxed{\text{A}}\wedge$NP$\boxed{\text{D}}$
VP/(NP$\boxed{\text{A}}\wedge$NP$\boxed{\text{D}}$)			
VP			

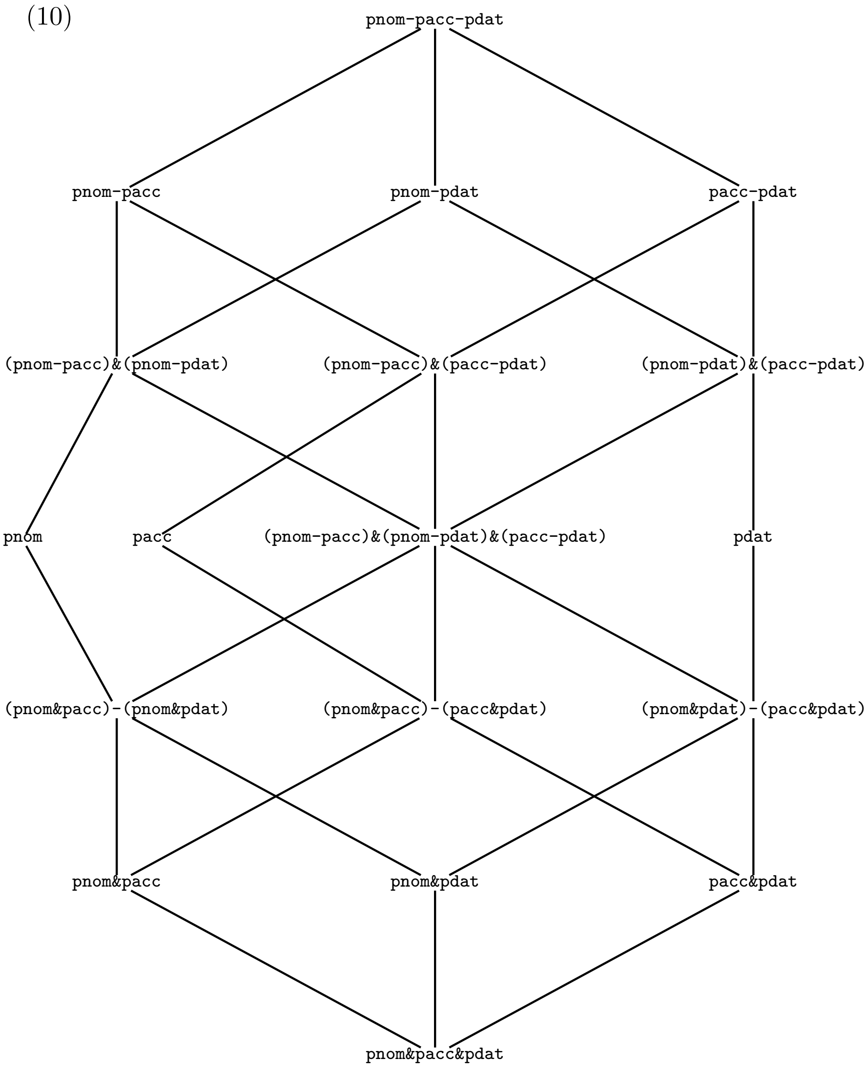
Several researchers have recently attempted to incorporate Bayer and Johnson's insights into HPSG. Levy (2001) augments the space of resolved feature structures in terms of objects he calls 'double-sets'.[2] These are organized into a lattice that is orthogonal to the familiar hierarchy of types assumed in HPSG work. Levy and Pollard (2001) adapt this idea in terms of boolean types (types built up via meet (&) and join ($-$) operations). On their proposal, three 'pure' case types such as *pnom*, *pacc*, and *pdat* give rise to 18 maximal types organized as follows:[3]

[2] The double-set lattice over the set $\{A,B\}$ is constructed from the following elements:

$\emptyset, \{\emptyset\}, \{\{A\}\}, \{\{B\}\}, \{\{A\},\{B\}\}, \{\{A,B\}\}$

[3] This is the Smyth powerlattice of the powerset (ordered by the subset relation) of a 3-element set, minus the top and bottom elements.

(10)

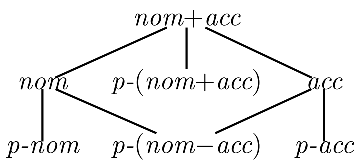


Levy and Pollard use this lattice to provide a space of maximal values that a coordinate NP's CASE value can be resolved to when the case of its conjuncts is not uniform. Join represents syncretization, and meet is coordination.[4]

[4]The hierarchies of Levy and Pollard are inverted (with respect to others discussed below, where (&) corresponds to syncretization and join (∨) corresponds to coordination).

In a related proposal, Daniels (2001) independently proposes a (semi-) lattice-based solution to some of the problems of feature neutrality and the coordination of unlikes. Daniels' solution, however, does not include a lattice structure independent of the familiar type hierarchy. Rather, he adapts ideas developed by Levine et al. (2001) for the analysis of apparent case discrepancies among parasitic gaps. Daniels posits hierarchies of CASE values that include the simplified example shown in (11):

(11)



This is a type hierarchy of the familiar kind, where only the leaf objects are maximal. What is perhaps unfamiliar here is the distinction between pure types (those begining with *p-*) and non-pure types. A linguistic object must be assigned a pure type. A syncretic expression is assigned a type constructed with −; a coordination of unlikes is assigned a type constructed with +. The case of a coordinate structure whose conjuncts are, for example, *p-nom* and *p-acc* is the pure type whose corresponding non-pure type is identical to, or a supertype of, the non-pure types *nom* and *acc*. Given the hierarchy in (11) then, the case of the coordinate structure is *p-*(*nom+acc*), a pure type that 'c-commands', as it were, the types of the conjuncts' cases in the type hierarchy.

Daniels' analysis posits hierarchies as rich as those offered by Levy and Pollard, but Daniels suggests that pieces of the hierarchy should be absent if the relevant syncretic forms are unattested in a given language. Otherwise, Daniels' proposal is in fact reducible to the one made by Levy and Pollard, as the latter authors note. To give the reader a feel for the complexity introduced into HPSG by these interrelated proposals, I will simply quote Levy and Pollard (2001: 225):

> So in a three-case system, this version of the Levine et al. hierarchy [upon which all of these proposals are based – IAS] would be isomorphic to the semilattice obtained by taking the powerset of a 7-element set and tossing out the empty set, giving a total of 127 nodes. In a four-case system, the Levine et al. hierarchy would have 32,767 nodes.

14.3 A Proposal

Reflecting on these recent attempts to incorporate into HPSG the insights of Bayer and Johnson, I am struck by two things: (1) the importance of Bayer's and Johnson's insights about how category resolution and coordination interact and (2) the complexity that is apparently required to reconcile that insight with the modeling assumptions of HPSG. In the remainder of this paper, I want to consider what may be a simpler way of incorporating into HPSG the insights of all the researchers whose work I have just reviewed. This involves making one small but significant modification to HPSG's modeling assumptions.

Let's begin with a simple observation. Though much has been made of the HPSG assumption that feature structures are 'totally well-typed' (bear a specification for all features that could be specified for that type of feature structure) and 'sort-resolved' (assigned a maximal type – one that has no subtypes),[5] the fact of the matter is that the sentence descriptions produced by HPSG grammars typically have only one intended feature structure that satisfy them. For example, the grammar rules, general principles, and lexical entries in Pollard and Sag 1994 are such that for any well-formed word string, there is one feature structure model satisfying each grammar-induced description of that string. That is, a given sentence may be ambiguous in virtue of lexical or structural ambiguity, but in that case the grammar will provide a distinct description for each alternative reading. Sentence models and sentence descriptions are in general isomorphic.

The one exception to this that comes to mind is Pollard and Sag's (1994) treatment of quantifier scope in terms of constraints on quantifier 'retrieval'. There the constraint defining the relation between the head daughter's STORE value, the mother's STORE value and the QUANTS values of mother and head daughter may be satisfied in more than one way if more than one quantifier is in the head daughter's STORE value. The result is a one-to-many relation between the grammar-induced sentence description and the feature structure models that satisfy that description.

But this is a treatment of quantification that has been called into question. In particular, the framework of Minimal Recursion Semantics (MRS)[6] eliminates the entire notion of 'storage' in favor of a system where the grammar characterizes signs with scope-neutral semantic structures. These unscoped CONTENT values are then related to fully resolved MRS structures by general principles that lie outside the system of constraints on well-formed feature structures provided by the grammar. If we adopt MRS, or some other approach to semantics that allows scope underspecification, then HPSG models and sentence descriptions will indeed be isomorphic.

[5]See Pollard and Sag (1994, chapter 1), King (1989, 1994), and Carpenter 1992.

[6]See Copestake et al. 1995, Copestake et al. 1999, and Copestake et al. 2001.

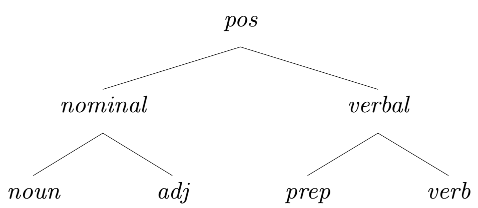
Given this observation, i.e. given the fact that the HPSG grammars we actually write provide fully determinate sentence descriptions, perhaps it is unnecessary to impose the additional requirement that structures be 'fully determinate', as reflected in the stipulation that feature structure models must be totally well-typed and sort-resolved. Here I will explore a simple modification of the standard HPSG modeling assumptions: abandoning the requirement that feature structures be sort-resolved. That is, feature structures will be specified for all features declared appropriate for their type but the values of those features need not be assigned leaf (maximal) types in the type hierarchy.

Following Levy, I will assume that verbs and other selectors impose a lower bound on the type of their arguments[7] and that the arguments themselves fix the relevant value (or else provide an upper bound). In the simplest cases, the specifications of a selector and those of the selected item coincide and the relevant value is uniquely determined. In other instances, those involving selectional underspecification, syncretic forms and the coordination of unlikes, the specifications of selector and the selected may diverge, as I will illustrate.

14.3.1 The Coordination of Unlikes

To get started, let us reconsider the coordination of expressions of unlike category and Jacobson's puzzle. I will assume, as have Levy, Pollard and Daniels, that the relevant part-of-speech distinctions are organized into a hierarchy like the one shown in (12):

(12)



The leaf types in (12) are maximal and the nonmaximal types *pos*, *nominal* and *verbal* correspond to the disjunctive types of the Categorial Grammar analyses. However, instead of allowing a fully expanded Boolean category space, as Bayer and Johnson do, I will follow Daniels in assuming that the only conjunctive and disjunctive types we have are those that are linguistically motivated. Conjunctive types are motivated by syncretism; disjunctive types by neutralization in coordinate

[7]More precisely: a greatest lower bound on the type of the value of some feature of each argument. The terminology may seem confusing, because the 'lower' types are displayed above the 'higher' types in the diagrams that appear below.

structures. I also follow Daniels (but not Levy and Pollard) in imposing only one hierarchy on types. That is, (12) is a type hierarchy of a familiar kind, where only the leaf types are maximal.

For the moment, let us assume that a verb like *elect* pins down its object's part-of-speech precisely, while verbs like *become* and *remain* specify a nonmaximal type that serves as a bound on their complement's part-of-speech:

(13) elect: $[\text{COMPS} = \langle[\text{HEAD} = \textit{noun}\ ]\rangle]$

(14) become, remain: $[\text{COMPS} = \langle[\text{HEAD} = \boxed{1}, \textit{nominal} \leq \boxed{1}\ ]\rangle]$

Note that I use 'less-than-or-equal-to' ($\leq$) to formulate bounding constraints, i.e. constraints that permit multiple resolutions. Since feature structures with nonmaximal values are now permitted, $[\text{HEAD} = \textit{nominal}]$ should be interpreted as fixing the HEAD value as (the nonmaximal type) *nominal*. In this type hierarchy, 'is less than' means 'is a supertype of'.

Lexical entries specify the appropriate maximal part-of-speech type in English, assuming English has no category-syncretic words. Hence the lexical entries in (13) and (14) are sufficient to account for standard simple data sets like the following:[8]

(15) a. They elected a Republican/*wealthy/*given a book...
b. Kim became/remained a Republican/wealthy/*given a book...

To deal with constituent coordination (other than NP coordination), consider the following simplified rule, which blends the approach of Shieber (1992) with that of Daniels (2001):

(16) General Coordination Rule ($\leq$-based)

$$\begin{bmatrix}\text{HEAD} = \boxed{0}\\ \text{VAL} = \boxed{V}\end{bmatrix} \rightarrow \begin{bmatrix}\text{HEAD} = \boxed{1}\\ \text{VAL} = \boxed{V}\end{bmatrix} \cdots \begin{bmatrix}\text{HEAD} = \boxed{n{-}1}\\ \text{VAL} = \boxed{V}\end{bmatrix} \text{CNJ} \begin{bmatrix}\text{HEAD} = \boxed{n}\\ \text{VAL} = \boxed{V}\end{bmatrix}$$

where $\boxed{0} \leq \boxed{1}$, ... , $\boxed{0} \leq \boxed{n{-}1}$, and $\boxed{0} \leq \boxed{n}$.

Here the HEAD value of the mother is constrained to be less-than-or-equal-to the HEAD value of each conjunct. However, again following Daniels, the VAL(ENCE) values of the conjuncts are identified with each other and with the mother's VAL value. Thus a stronger condition is placed on the features used for argument selection.

[8]For convenience, I'm ignoring the fact that *remain*, but not *become*, is compatible with a PP complement.

We may now illustrate how the coordination of unlikes is analyzed. The constraints included in the lexical entries given above interact with the ≤-based formulation of the Coordination Rule given in (16) to allow examples like (17). The HEAD value of each relevant element is uniquely determined, as illustrated:

(17) a. become: [COMPS ⟨[HEAD = $\boxed{1}$, *nominal* ≤ $\boxed{1}$]⟩]
b. wealthy: [HEAD = *adj*]
c. (a) Republican: [HEAD = *noun*]
d. wealthy and a Republican: [HEAD = *nominal*]

This analysis relies crucially on the assumption that the HEAD value of the coordinate phrase is a feature structure assigned to the nonmaximal type *nominal*, i.e. it relies on the assumption that feature structures need not be sort-resolved.

And because VAL values are identified in coordination, the COMPS value of a verbal coordination like *grew and remained* will be subject to all the constraints imposed by the conjuncts. Since the constraints of *grew* are more specific than those of *remain*, it follows that *grew and remained* must obey the more specific constraints, as shown in (18):

(18) a. grew: [COMPS = ⟨[HEAD = *adj*]⟩]
b. remained: [COMPS = ⟨[HEAD = $\boxed{1}$, *nominal* ≤ $\boxed{1}$]⟩]
c. grew and remained: [COMPS = ⟨[HEAD = *adj*]⟩]
d. grew and remained wealthy
e. *grew and remained a Republican
f. *grew and remained wealthy and a Republican

This provides an HPSG solution to Jacobson's puzzle with a less complicated hierarchy than those assumed by Levy, Pollard, or Daniels. Moreover, the relevant constraints are all stated simply in terms of the notion of '≤'.[9] Finally, it should be noted that there is no spurious ambiguity in this analysis: the constraints imposed by the grammar are such that each example we have considered has at most one feature structure model.

There is a slightly different approach that we should also consider. Suppose that the lexical entries for nouns and arguments did not fix

[9] We haven't really considered all the relevant data yet. For example, each different kind of unlike category coordination would motivate positing a new supertype in my analysis, since there must be some nonmaximal type for the HEAD value of the coordinate structure to resolve to. Even if the hierarchy in (12) must be further complicated, however, it will still have significantly fewer types than the alternative analyses just mentioned, and will use simpler constraints.

the type of the HEAD value, but rather put an upper bound on it, as shown in (19):

(19) a. wealthy: $[\text{HEAD} = \boxed{1}\ ,\ \boxed{1} \leq adj\ ]$
b. (a) Republican: $[\text{HEAD} = \boxed{1}\ ,\ \boxed{1} \leq noun]$

Leaving all other aspects of our analysis unchanged, this would still provide a solution to the problem of unlike category coordination. The grammar would allow exactly one analysis for *became wealthy and a Republican*: the type of the HEAD value of the coordinate complement would be *nominal*. Note that this alternative proposal would allow us to modify our Coordination Rule so that it imposes the stronger condition that the HEAD value of the conjuncts and their mother be identical:

(20) General Coordination Rule (=-based)

$$\begin{bmatrix}\text{HEAD} = \boxed{0}\\ \text{VAL} = \boxed{1}\end{bmatrix} \rightarrow \begin{bmatrix}\text{HEAD} = \boxed{0}\\ \text{VAL} = \boxed{1}\end{bmatrix} \ldots \begin{bmatrix}\text{HEAD} = \boxed{0}\\ \text{VAL} = \boxed{1}\end{bmatrix} \text{CNJ} \begin{bmatrix}\text{HEAD} = \boxed{0}\\ \text{VAL} = \boxed{1}\end{bmatrix}$$

I believe there is no English evidence to distinguish among the three proposals just outlined: (1) lexical entries (17b,c) with the $\leq$ version of the Coordination Rule, (2) lexical entries (19a,b) with the $\leq$ version of the Coordination Rule, and (3) lexical entries (19a,b) with the = version of the Coordination Rule. All three analyses provide an account of unlike category coordination that solves Jacobson's puzzle without creating spurious ambiguity, and hence appear to be empirically indistinguishable.[10]

[10] As Roger Levy and Adam Przepiórkowski both remind me, there is a further potential problem illustrated by Polish examples like the following (Przepiórkowski 1998, ex. (5.265)):

(i) Jana dziwi, [[że Maria wybiera Piotra], i [jej brak
John-ACC suprise Comp Mary-NOM chooses Peter-ACC and her lack-NOM
gustu]].
good taste-GEN

'John is surprised that Mary chooses Peter and by her lack of good taste.'

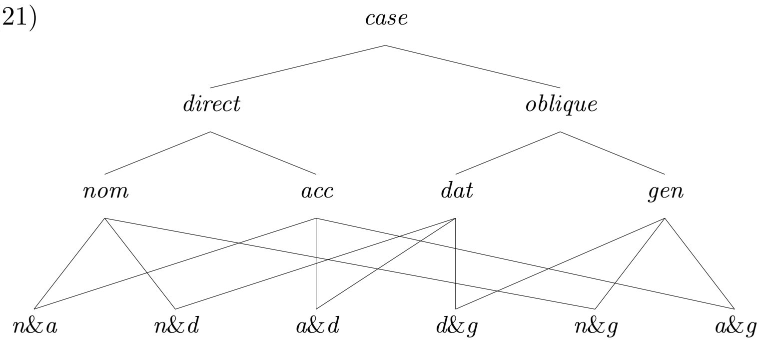
If we assume that assignment of case (nominative in the case of (i)) must be preserved in cross-categorial coordination, then it appears that case distinctions must somehow be reflected in the hieararchy of categories. Under this assumption, (i) presents the same dilemma for Levy's, Levy and Pollard's and Daniels' analyses.

I am not at all sure that case has to be transmitted across unlike category coordination, but let us suppose it does. A simple solution to this problem involves letting CPs (and thus, complementizers) bear CASE specifications, as, for example, in Sag et al. (to appear). This treatment of CPs is independently motivated by the fact that case-assigning verbs often allow CPs in argument positions (e.g. *He bothers me*/*That Sandy left bothers me*). An alternative approach might treat *nom* as a default CASE value, leaving the mother of the coordinate structure in (i) unpecified for the feature CASE.

14.3.2 Case Neutralization in German

Now let us reconsider German case neutralization. Here, following Levy, Pollard, and Daniels (and Bayer and Johnson), we may posit conjunctive types to allow for the possibility of syncretic forms. The case system of German can then be based on the following hierarchy of types, where *direct* and *oblique* are familiar disjunctive types:

(21)



Note that conjunctive case types are posited only if German contains some syncretic word that actually is neutral over the relevant cases in coordination.

Because it allows for a simpler formulation of relevant constraints,[11] I will assume here that German makes use of the =-based general coordination rule given in (20) above, rather than its $\leq$ counterpart in (16). The analysis proceeds as follows. First, the verbs *findet* and *hilft* constrain the case of their object NP in the following ways:

(22) findet: $[\text{COMPS} = \langle[\text{CASE} = \boxed{1}\ ,\ acc \leq \boxed{1}\ ]\rangle]$

(23) hilft: $[\text{COMPS} = \langle[\text{CASE} = \boxed{1}\ ,\ dat \leq \boxed{1}\ ]\rangle]$

Then, since we adopt the =-based formulation of general coordination, we posit a lexical entry for *Männer* that contains the constraint shown in (24a):

(24) a. Männer: $[\text{HEAD} = [\text{CASE} = \boxed{1}\ ,\ \neg(dat \leq \boxed{1}\ )\ ]]$
b. Kindern: $[\text{HEAD} = [\text{CASE} = \boxed{1}\ ,\ \boxed{1} \leq\ dat\ ]]$

Similarly, the dative noun *Kindern* has a lexical entry that includes the specification shown in (24b), which makes this word incompatible with

[11]The identity-based coordination rule might be simplified further, say, by identifying the entire CATEGORY, LOCAL, or SYNTAX value of each conjunct with that of the mother. This simplification appears unavailable within the $\leq$-based alternative.

all nondative contexts. These assumptions suffice to account for simple case restrictions, as shown in (25):

(25) a. findet Männer: [CASE = *acc*]
b. *findet Kindern: [CASE = ??]
c. hilft Kindern: [CASE = *dat*]
d. *hilft Männer: [CASE = ??]

Note that on this analysis, if a given word's case is unambiguous, then its lexical entry provides an upper bound on the relevant CASE value, as illustrated in (26):

(26) a. ich: [HEAD = [CASE = $\boxed{1}$, $\boxed{1} \leq nom$]]
b. dich: [HEAD = [CASE = $\boxed{1}$, $\boxed{1} \leq acc$]]
c. des: [HEAD = [CASE = $\boxed{1}$, $\boxed{1} \leq gen$]]

This will play a key role in the treatment of NP coordination sketched below, which must ensure, for example, that nouns like (26a) cannot be coordinated with nouns like (26b).

A syncretic noun like *Frauen* can be resolved to any case. Thus its lexical entry needn't mention case (assuming that the grammar signature ensures that any CASE value is greater-than-or-equal-to *case*). This means that *Frauen* will be allowed as an object in contexts that require conjunctive CASE values. For example, if we coordinate *findet* and *hilft*, the resulting verb must satisfy the valence requirements of both verbs. This is impossible with nouns whose CASE value is incompatible with *a&d* (e.g. those in (24)), but it is possible with *Frauen*, as illustrated in (27):

(27) a. findet und hilft: [COMPS = ⟨[CASE = *a&d*]⟩]
b. findet und hilft Frauen: [CASE = *a&d*]
c. *findet und hilft Männer: [CASE = ??]
d. *findet und hilft Kindern: [CASE = ??]

I'll turn to nominal coordination in a moment. But first, let's consider the issue of 'spurious' ambiguity. The analysis I have just sketched in fact allows four values for the case of *Frauen* when it occurs as the object of *findet* or *hilft*:

(28) a. findet Frauen: [CASE = $\boxed{1}$, $\boxed{1} \in \{acc,\ n\&a,\ a\&d,\ a\&g\}$]
b. hilft Frauen: [CASE = $\boxed{1}$, $\boxed{1} \in \{dat,\ n\&d,\ a\&d,\ d\&g\}$]

It also allows three values for the case of *Männer* occurring as the object of *findet*:

(29) findet Männer: [HEAD = [CASE = $\boxed{1}$, $\boxed{1} \in \{acc,\ n\&a,\ a\&g\ \}$]]

My analysis thus seems to introduce multiple analyses that correspond to no linguistic ambiguity, a fact that might be construed as an argument against it.

However, there are at least two responses that could be made to this objection. First, as Ken Shan has suggested to me (personal communication), one might simply revise the way that linguistic ambiguity is defined so that families like those in (28) and (29) (feature structures that differ merely with respect to contiguous types) constitute an equivalence class. This proposal could also be thought of as letting a 'region' of the type hierarchy count as a single linguistic object. I'm not sure what undesirable consequences (if any) Shan's proposal might have; but it has a certain resemblance to proposals within Categorial Grammar to let semantically equivalent analyses count as linguistically nondistinctive.

An alternative solution involves altering the way 'root' signs are defined along the following lines:

(30) A feature structure F corresponds to a 'stand-alone' utterance with respect to a grammar $\mathcal{G}$ just in case F satisfies:

1. all constraints of $\mathcal{G}$, and
2. $$\begin{bmatrix} \textit{sign} \\ \text{LOC} = \begin{bmatrix} \text{CAT} = \begin{bmatrix} \text{HEAD} = \begin{bmatrix} \text{VFORM} = \textit{fin} \end{bmatrix} \\ \text{SUBJ} = \langle\ \rangle \\ \text{COMPS} = \langle\ \rangle \end{bmatrix} \end{bmatrix} \\ \text{SLASH} = \{\ \} \end{bmatrix},$$

and there is no F' more general than F that also satisfies 1 and 2.

Nothing here hinges on the specifics of the root condition given in (30)2.[12] The effect of the definition in (30) is to restrict the root-level signs defined by a grammar to the most general satisfiers of the grammar's constraints. As a result, all but the first feature structure schematized in (28a), (28b), or (29) would be eliminated from consideration. And with these feature structures out of the picture, the spurious ambiguity problem would be eliminated.

This approach to the spurious ambiguity problem will seem more satisfying to most linguists, I suspect. However, since it involves consider-

[12] But for a defense of the idea that S[*fin*] should be the category of utterances consisting of only elliptical XP fragments, see Ginzburg and Sag 2000.

ing multiple feature structures in order to ascertain well-formedness of any single feature structure, it might be objected that we have pushed HPSG outside the realm of 'model-theoretic grammar' (in the sense of Pullum and Scholz 2001). I don't think this objection cuts very deep, however. Because we have abandoned only the condition that feature structures must be totally well-typed, the determination of 'most general' is entirely local. That is, for any given feature structure, we need only consider a small space of alternative types in order to determine whether the assigned type is the most general one compatible with the relevant constraints. And this is all that needs to be considered in order to determine well-formedness. Thus the notion of 'most general satisfier' of a set of constraints that I am appealing to here seems unproblematic.[13]

Finally, let us consider NP coordination in German and English, which I will assume can be analyzed via a rule like the following:

(31) NP Coordination Rule:

$$\overset{\text{NP}}{\begin{bmatrix}\text{NUM} = pl\\ \text{PER} = \boxed{0}\\ \text{CAT} = \boxed{C}\end{bmatrix}} \rightarrow \overset{\text{NP}}{\begin{bmatrix}\text{PER} = \boxed{1}\\ \text{CAT} = \boxed{C}\end{bmatrix}} \ldots \overset{\text{NP}}{\begin{bmatrix}\text{PER} = \boxed{n{-}1}\\ \text{CAT} = \boxed{C}\end{bmatrix}} \text{ CNJ } \overset{\text{NP}}{\begin{bmatrix}\text{PER} = \boxed{n}\\ \text{CAT} = \boxed{C}\end{bmatrix}}$$

where $\boxed{1} \leq \boxed{0}$, ... , $\boxed{n{-}1} \leq \boxed{0}$, and $\boxed{n} \leq \boxed{0}$.

This rule stipulates that all coordinate NPs are plural.[14] It also requires that NP conjuncts share their CAT value and that a coordinate NP's PER(SON) value be determined by the following hierarchy:

(32)

$$\begin{array}{c} \textit{3rd} \\ | \\ \textit{2nd} \\ | \\ \textit{1st} \end{array}$$

This proposal builds into the analysis of NP coordination a version of the treatment of person/coordination interactions that is originally

[13] Note also that there need not be a unique most general satisfier of any particular set of constraints. If we assume that any two compatible types have a unique $\leq$ bound, however, then uniqueness can be guaranteed. There are further issues here having to do with set-valued features, but these are beyond the scope of the present paper.

[14] This is a simplification in that there is unclarity about the NUMBER value of NPs coordinated with *or*. For some discussion, see Morgan (1972, 1984).

cast in set-theoretic terms by Sag et al. (1985). However, my analysis is in one respect more like that of Dalrymple and Kaplan (2000), in that it is based on the values of the feature PER (rather than on sets of feature-value pairs, as in GPSG). Set-theoretic relations are here replaced by the ordering of the type hierarchy,[15] which makes it possible to capture Sag et al's generalization that 'the PERSON value of a coordinate NP is the minimum of the persons of the conjuncts'. 'Minimum' is here interpreted as 'most specific'.

The CAT identities specified in (31) ensure that the CASE values of all NP conjuncts must be identical. This generalization rings true, even though the particular case one finds in pronominal conjuncts is often not the expected one:

(33) a. He and I left.
b.%Him and me left.
c.%They invited Kim and I.

(34) a. *Him and I left
b. *I and him left.
c. *They invited him and I.
d. *They invited I and him.

These judgments reflect only a cursory exploration of dialects where (33b) or (33c) are grammatical. Clearly, a more thorough investigation of these data is called for.

In German, we find NP structures like the following, where the CASE value of the coordinate structure is just as resolved as that of the most specific conjunct:

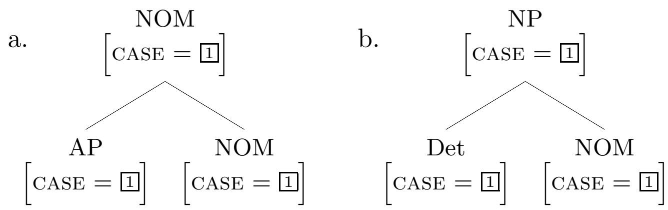
(35) Kindern und Frauen: [CASE = $\boxed{1}$, $\boxed{1} \leq$ *dat*]

In addition, identity of CASE value will be imposed in other NP-internal configurations, e.g. the following:

[15] Given the following correspondence, it is clear that the $<$ relation in my analysis corresponds to the subset relation, as used in Dalrymple and Kaplan's (D & K's) analysis of English (as pointed out in passing by Levy and Pollard (2001)):

TRADITIONAL CATEGORY	D & K's PER VALUE	MY PER VALUE
3rd person	$\emptyset$	*3rd*
2nd person	$\{\mathbf{H}\}$	*2nd*
1st person	$\{\mathbf{S}, \mathbf{H}\}$	*1st*

(36)



This will provide a correct account of the following contrasts (among others), which are discussed by Levy (2001):

(37) a. Er findet (die/*den) Frauen.
He finds (the-ACC/*the-DAT) women.

b. Er hilft (*die/den) Frauen.
He helps (*the-ACC/the-DAT) women.

c. Er findet und hilft (*die/*den) Frauen.
OBJ.ACC OBJ.DAT *ACC/*DAT
He finds and helps *the women

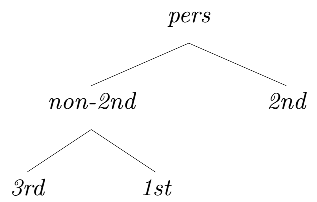
In (37a), the case of *den*, which is identified with that of the entire direct object NP, must be less than or equal to *dat*, but no such case is greater than or equal to *acc*, as required by *findet*. In (37b), the case of *die* must be less than or equal to *acc*, which is incompatible with being greater than or equal to *dat*, as required by the constraint introduced by the verb *hilft* as a bound on the object's case. Finally, in (37c), the case of the object of *findet und hilft* can only be *a&d*, which is incompatible with the case of either *die* or *den*. Since the determiner's case must be the object's case, neither of these options is well-formed. The contrasts in (37) are thus predicted to the letter.

That said, I have to confess that the person analysis just offered seems to be at odds with one of the standard examples in the coordination/syncretism literature, namely the following:

(38) ...weil wir das Haus und die Müllers den Garten
because we the house and the Müllers the garden
kaufen.
buy-1pl/3pl
...because we buy the house and the Müllers, the garden

I'm really not sure how to analyze right node raising examples like this, but if the subject requirements of the verb *kaufen* must somehow be satisfied by both *wir* (1st person) and *die Müllers* (3rd person), then it would appear that there must be a person type that is neutral to (a supertype of) these two possibilities, as sketched in (39):

(39)

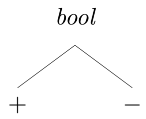


Needless to say, this person hierarchy is inconsistent with the one in (32) above. At present, I can only flag this as an unresolved issue.

14.3.3 English Auxiliaries

The feature AUX has long been problematic for HPSG analyses of English. VP conjuncts need not agree on AUX values (though they must agree on values of other HEAD features, e.g. VFORM), as illustrated in (40)–(41):

(40)



(41) a. likes bagels: [AUX $= -$]
b. is happy: [AUX $= +$]
c. Kim [likes bagels and is happy]: [AUX = ??]

The present framework provides an immediate solution to this dilemma. The AUX values are now lexically constrained as in (42a,b):

(42) a. likes (bagels): [AUX $= \boxed{1}$, $\boxed{1} \leq -$]
b. is (happy): [AUX $= \boxed{1}$, $\boxed{1} \leq +$]
c. Kim [likes bagels and is happy]: [AUX = *bool*]

And hence, as long as no more specific constraint is imposed, *bool* may serve as the AUX value of the coordinate VP, as shown in (42c). This provides an account of why such discrepancies do not give rise to ungrammaticality.

14.3.4 English Right Node Raising

Data like the following, noted in section 1, are discussed by Pullum and Zwicky (1986):

(43) I certainly will, and you already have, {*clarify/*clarified the situation}
{set the record straight} with respect to the budget.

Contrasts like these should lend themselves to a solution similar to those already presented. Again, there is uncertainty about how to analyze the right node raising construction, but if the correct analysis involved a feature whose value would have to satisfy the constraints imposed by the verbs governing the VP gaps as well as those specified in the lexical entry of the head of the right-raised VP, then the solution should be exactly like the others we have seen. The hierarchy of VFORM values, whatever it turns out to be, must include the subhierarchy shown in (44):

(44)

$$\begin{tikzcd}[column sep=small] & \mathit{vform} \arrow[dl, no head] \arrow[dr, no head] & \\ \mathit{pfp} \arrow[dr, no head] & & \mathit{bse} \arrow[dl, no head] \\ & \mathit{pfp}\&\mathit{bse} & \end{tikzcd}$$

The lexical entry for *set* includes the constraint shown in (45a):

(45) a. set: [HEAD = [VFORM = $\boxed{1}$, $\boxed{1} \leq$ *pfp*&*bse*]]
b. clarify: [HEAD = [VFORM = $\boxed{1}$, $\boxed{1} \leq$ *bse*]]
c. clarified: [HEAD = [VFORM = $\boxed{1}$, $\boxed{1} \leq$ *pfp*]]

Hence the VFORM value of *set* can be resolved to the conjunctive type *pfp*&*bse*, *set* can appear in right-raised contexts like (43). However, because both *clarify* and *clarified* have lexical entries that fix the VFORM value as indicated in (45b-c), neither can satisfy the constraints imposed by both *will* and *have* simultaneously:

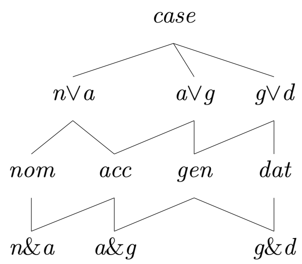
(46) a. will: [COMPS = ⟨ [VFORM = $\boxed{1}$, *bse* $\leq \boxed{1}$] ⟩]
b. have: [COMPS = ⟨ [VFORM = $\boxed{1}$, *psp* $\leq \boxed{1}$] ⟩]

This is what would be required in order for them to appear in these contexts.

14.3.5 Polish Case

The syncretic nouns of Polish would seem to motivate a hierarchy of case like the following:

(47)



The data relevant to establishing the conjunctive types in this hierarchy include constrasts like the following, noted in section 1:

(48) Kogo/*Co Janek lubi a Jerzy
(ACC/GEN)/*(NOM/ACC) OBJ.ACC
who John likes and George
nienawidzi?
OBJ.GEN
hates
"Who/*What does John like and George hate?"
(Polish: Dyła 1984)

That is, the syncretic nouns make reference to the conjunctive types in lexical entries like the following:

(49) a. kogo: [HEAD = [CASE = $\boxed{1}$, $\boxed{1}$ $\leq$ *a&g*]]
b. co: [HEAD = [CASE = $\boxed{1}$, $\boxed{1}$ $\leq$ *n&a*]]

Polish verbs, like those considered above, place a bound on the CASE value of their object − *acc* and *gen* in the case of the verbs in (48). Thus the clauses that are coordinated in (48) are specified as shown in (50a,b):

(50) a. Janek lubi: [SLASH = {[CASE = $\boxed{1}$, *acc* $\leq$ $\boxed{1}$]}]
b. Jerzy nienawidzi: [SLASH = {[CASE = $\boxed{1}$, *gen* $\leq$ $\boxed{1}$]}]
c. [[Janek lubi] a [Jerzy nienawidzi]]:
[SLASH = {[CASE = *a&g*]}]

Hence the coordinate clause bears the SLASH specification shown in (50c). As a result, the fronted element in such examples must be consistent with [CASE = *a&g*], i.e. it must be (or be headed by) a noun like *kogo*, not by *co*, and not by any nonsyncretic noun.

Finally, let us reconsider Przepiórkowski's (1999) example that was cited in section 1:

(51) Dajcie [wina i całą świnię]!
give wine.GEN and whole.ACC pig.ACC
"Serve some wine and a whole pig!"

My take on this example may be overly simplistic, but it seems that one can use the disjunctive type $a \vee g$ to let the verb *dajcie* place the appropriate lower bound on its object, as shown in (52a):

(52) a. dajcie: [COMPS = ⟨[CASE = $\boxed{1}$, $a \vee g \leq \boxed{1}$]⟩]
b. wina: [CASE = *gen*]
c. świnię: [CASE = *acc*]
d. dajcie [wina i całą świnię]: [CASE = $a \vee g$]

Assuming that the relevant nouns are specified as in (52b,c), then the NP coordination in (52d) is correctly analyzed, as long as Polish also uses the NP coordination rule proposed earlier.

Finally, there is further data, discussed by Levy (2001), that is also properly accounted for by this analysis:

(53) a. *? [Maria kocha a Ewa nienawidzi]
Maria loves.OBJ-ACC but Ewa hates.OBJ-GEN
tego mężczyzny.
this.ACC/GEN man.ACC

b. [Maria kocha a Ewa nienawidzi] tego
Maria loves but Ewa hates this.ACC/GEN
faceta.
guy.ACC/GEN
(Przepiórkowski, personal communication to Levy)

(54) *Včera vec' den' on proždal [svoej podrugu
yesterday all day he awaited self's-GEN girlfriend-ACC
Irinu] i [zvonka [ot svoego brata Grigorija]].
Irina.ACC and call-GEN [from self's brother Gregory]

If we assume that the coordinate clauses in (53a,b) work in essentially the same way as leftward extraction examples like (48), i.e. via inheritance of SLASH specifications, then the CASE value of the right-raised NP must be $a \& g$. The determiner *tego* can resolve to this value, as can *faceta*, but the nonsyncretic noun *mężczyzny* cannot (it is upper-bounded by *acc*). Therefore, because the CASE value of the right-shifted NP must be the same as that of its head noun and that of its determiner (see (36) above), the contrast between (53a,b) is correctly predicted. Similarly, the Russian example (54) is ruled out, because the left conjunct's modifier must have the same case as its head, which it cannot.

Though the coordinate NP can be neutral with respect to *acc* and *gen*, each NP conjunct must be internally case-consistent.

14.4 Conclusion

Following foundational work by King (1989, 1994), Pollard and Sag (1994) and others working in HPSG have made the assumption that feature structures must be 'fully specified'. This notion has been interpreted as meaning 'totally well-typed' (bear a specification for all features that could be specified for that type of feature structure) and 'sort-resolved' (assigned to a maximal type). Ingria's (1990) much-cited paper (and Zaenen and Karttunen's (1984) important precursor) discussed data from various languages that pose a serious challenge for these assumptions. These problems and others were addressed in work by Bayer and Johnson (1995) and Bayer (1996), who propose an analysis in terms of Type Logical (Categorial) Grammar.

A number of recent attempts have been made (Levy 2001, Levy and Pollard 2001, and Daniels 2001) to integrate Bayer and Johnson's insights into HPSG accounts of the troublesome data involving coordination of unlikes, feature neutralization, case syncretism, and related issues. These proposals, however, have imposed new hierarchies on maximal types or else have introduced considerable complexity into existing type hierarchies, a complexity that I have tried to eliminate in this paper.

Eliminating the requirement that feature structures be assigned maximal types, I have suggested that it is possible to simplify these analyses, eliminating the need for new hierarchies, while nonetheless incorporating the insights of the Type Logical analyses and the solutions they provide to problems noted by Zaenen and Karttunen, Pullum and Zwicky, Ingria, and Jacobson. Of course, it may prove to be desirable to make a more radical departure from King's foundational assumptions, by introducing partiality more generally. And this may well be possible (eliminating the totally well-typed requirement as well, for example), for, as I have noted, the constraints induced by an HPSG grammar arguably uniquely determine a feature structure model for each desired sentence type without additional foundational assumptions. I leave open the possibility of further modifications along these lines.[16] My goal here has been to explore a minimal modification of familiar theoretical foundations, which seems to make available straightforward

[16] It is also possible that some of the examples discussed here should be analyzed as discontinuous dependencies, as Jim Blevins has suggested to me on many occasions. On such an approach, examples like (1) would involve VP coordination, where only the first conjunct's lexical head is phonologically realized.

accounts of the diverse phenomena I have surveyed.

Acknowledgments

This paper was written while I was a fellow at the Center for Advanced Study in the Behavioral Sciences in Stanford, CA. I gratefully acknowledge the support of a grant (# 2000-5633) to CASBS from The William and Flora Hewlett Foundation. I am particularly indebted to Jim Blevins, with whom I've been discussing coordination and related issues since 1994, and to Berthold Crysmann, Roger Levy and Adam Przepiórkowski, who gave me detailed comments on an earlier draft. Thanks also to those who participated in my 2002 Universal Grammar seminar at Stanford. In addition, I would like to thank the following people for valuable discussion and/or helpful suggestions: Anne Abeillé, John Beavers, Emily Bender, Bob Borsley, Luis Casillas, Mike Daniels, Dan Flickinger, Mark Johnson, Ron Kaplan, Lauri Karttunen, Bob Levine, Rob Malouf, Carl Pollard, Chris Potts, Ken Shan, Stuart Shieber, Mark Steedman, Aline Villavicencio, Tom Wasow, Joel Wallenberg, and Arnold Zwicky. I see no reason to attribute to any of these people agreement (or disagreement) with the ideas I have explored here.

References

Bayer, Samuel. 1996. The coordination of unlike categories. *Language* 72: 579–616.

Bayer, Samuel and Mark Johnson. 1995. Features and agreement. *Proceedings of the 33rd Annual Meeting of the Association for Computational Linguistics* Pp. 70–76.

Carpenter, Bob. 1992. *The Logic of Typed Feature Structures*. Cambridge Tracts in Theoretical Computer Science 32. Cambridge University Press.

Copestake, Ann, Daniel Flickinger, Ivan A. Sag, and Carl Pollard. 1999. Minimal Recursion Semantics: an introduction. Manuscript, Stanford University: CSLI. (to be available in revised form in January, 2003).

Copestake, Ann, Daniel Flickinger, Robert Malouf, Susanne Riehemann, and Ivan A. Sag. 1995. Translation using Minimal Recursion Semantics. In *Proceedings of The Sixth International Conference on Theoretical and Methodological Issues in Machine Translation*. Leuven. Pp 15–32.

Copestake, Ann, Alex Lascarides, and Dan Flickinger. 2001. 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.

Dalrymple, Mary, and Ronald M. Kaplan. 2000. Feature indeterminacy and feature resolution. *Language* 76(4): 759-798.

Daniels, Michael. 2001. On a type-based analysis of feature neutrality and the coordination of unlikes. Paper presented at the HPSG-2001 Conference – Trondheim, Norway. Available at: http://cslipublications.stanford.edu/HPSG/2/hpsg01-toc.html

Dyła, Stephan. 1984. Across-the-board dependencies and case in Polish. *Linguistic Inquiry* 15: 701-705.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum, and Ivan A. Sag. 1985. *Generalized Phrase Structure Grammar*. Oxford: Basil Blackwell's and Cambridge, MA: Harvard University Press.

Ginzburg, Jonathan and Ivan A. Sag. 2000. *Interrogative Investigations: the form, meaning, and use of English Interrogatives.* Stanford: CSLI Publications. [Distributed by University of Chicago Press]

Ingria, Robert. 1990. The limits of unification. *Proceedings of the 28th ACL*. Pp. 194-204.

Jacobson, Pauline. 1987. Review of Gerald Gazdar, Ewan Klein, Geoffrey K. Pullum, and Ivan A. Sag, 1985: *Generalized Phrase Structure Grammar. Language.*

King, Paul John. 1989. *A Logical Formalism for Head-Driven Phrase Structure Grammar*. Ph.D. dissertation, University of Manchester.

King, Paul John. 1994. An expanded logical formalism for Head-driven Phrase Structure Grammar. *Arbeitspapiere des SFB 340* Nr. 59. Universität Tübingen.

Levine, Robert, Thomas Hukari, and Michael Calcagno. 2001. Parasitic gaps in English: some overlooked cases and their theoretical implications. In Peter Culicover and Paul Postal (eds.), *Parasitic Gaps*. Cambridge, MA: MIT Press. Pp. 181–222.

Levy, Roger. 2001. Feature indeterminacy and the coordination of unlikes in a totally well-typed HPSG. Draft of March 19, 2001. Available at: http://www.stanford.edu/~rog/papers/feature-indet.ps

Levy, Roger, and Carl Pollard. 2001. Coordination and neutralization in HPSG. Paper presented at the HPSG-2001 Conference – Trondheim, Norway. Available at: http://cslipublications.stanford.edu/HPSG/2/hpsg01-toc.html

Morgan, Jerry L. 1972. Verb agreement as a rule of English. *Proceedings of the 8th Regional Meeting of the Chicago Linguistic Society.* Chicago: CLS. Pp. 278-86.

Jerry L. Morgan. 1984. Some problems of determination in English number agreement. *ESCOL 84*. Somerville, MA: Cascadilla Press.

Pollard, Carl, and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press and Stanford: CSLI Publications.

Przepiórkowski, Adam. 1999. *Case Assignment and the Complement-Adjunct Dichotomy: A non-Configurational Constraint-Based Approach*. Ph.D. Dissertation, University of Tübingen.

Pullum, Geoffrey K., and Barbara C. Scholz. 2001. On the distinction between model-theoretic and generative-enumerative syntactic frameworks. In *Logical Aspects of Computational Linguistics: 4th International Conference* (Lecture Notes in Artificial Intelligence, 2099), ed. by Philippe de Groote, Glyn Morrill, and Christian Retoré, 17-43. Berlin: Springer Verlag.

Pullum, Geoffrey, and Arnold Zwicky. 1986. Phonological resolution of syntactic feature conflict. *Language* 62: 751-773.

Sag, Ivan A., Gerald Gazdar, Thomas Wasow and Steven Weisler. 1985. Coordination and how to distinguish categories. *Natural Language and Linguistic Theory* 3: 117–171. [Originally appeared as a CSLI Technical Report – March, 1984]

Sag, Ivan A., Thomas Wasow and Emily Bender. to appear. *Syntactic Theory – A formal introduction*, 2nd edition. Stanford: CSLI Publications. [To be distributed by University of Chicago Press].

Shieber, Stuart. 1992. *Constraint-Based Grammar Formalisms.* Cambridge, MA: MIT Press.

Zaenen, Annie and Lauri Karttunen. 1984. Morphological non-distinct-iveness and coordination. *Eastern States Conference on Linguistics* 1: 309–320.