Three improvements to the HPSG model theory

Adam Przepiórkowski

University of Warsaw / ICS Polish Academy of Sciences / University of Oxford

Proceedings of the 28th International Conference on
Head-Driven Phrase Structure Grammar

Online (Frankfurt/Main)

Stefan Müller, Nurit Melnik (Editors)

2021

Frankfurt/Main: University Library

pages 165–185

Keywords: RSRL, exhaustive models, rooted models, Höhle's Problem, unlike category coordination, HPSG

Przepiórkowski, Adam. 2021. Three improvements to the HPSG model theory. In Stefan Müller & Nurit Melnik (eds.), *Proceedings of the 28th International Conference on Head-Driven Phrase Structure Grammar, Online (Frankfurt/Main)*, 165–185. Frankfurt/Main: University Library. DOI: 10.21248/hpsg.2021.9.

Abstract

The aim of this paper is to propose three improvements to the HPSG model theory. The first is a solution to certain formal problems identified in Richter 2007. These problems are solved if HPSG models are rooted models of utterances and not exhaustive models of languages, as currently assumed. The proposed solution is compatible with all existing views on the nature of objects inhabiting models. The second improvement is a solution to "Höhle's Problem", i.e., the problem of massive spurious ambiguities in models of utterances. The third is a formalisation of Yatabe's (2004) analysis of the coordination of unlike categories, one that requires a second-order extension of the language for stating HPSG grammars.

1 Introduction

HPSG is unique amongst contemporary linguistic frameworks in having a well-developed model theory, most comprehensively presented in Richter 2004 (see Richter 2021 for an overview). Nevertheless, there are a number of problems that this model theory faces and there are some linguistic analyses that seem to call for an extension of that standard model theory.

In this paper, I propose three orthogonal improvements to the HPSG model theory of Richter 2004. Two of them address problems which are known and to some extent have been dealt with in the past. The first improvement, presented in Section 2 and Appendix A, deals with problems identified in Richter 2007, namely, the problems of HPSG models containing structures which are not linguistically motivated. The improvement consists in giving up the idea that models are exhaustive and allowing for rooted models.

The second proposal, presented in Section 3 and Appendix B, is not exactly an improvement of the model theory, but rather of the underyling grammars. It aims to solve what is sometimes (e.g., in Pollard 2001, 2014: 113) called "Höhle's Problem", i.e., the problem of massive spurious ambiguities in HPSG models, which are not intended – and not even suspected – by linguists writing their grammars. The solution consists in proposing certain constraints, assumed to be universal (i.e., parts of all grammars), which make sure that structures which look the same are token-identical. Such constraints have been proposed in the past, and what is new in the current proposal is a technique of exempting certain structures – especially, values of INDEX – from the scope of such constraints.

The third improvement extends the language in which HPSG theories are formulated in such a way that second-order statements are possible, i.e., in a way that makes it possible to refer not only to objects in the model but

[†]I am grateful for comments from Frank Richter and Manfred Sailer, as well as HPSG 2021 reviewers and the audiences of HPSG 2021 and the *Oberseminar Syntax and Semantics 2021* in Frankfurt. Needless to say, all remaining errors are mine alone.

also to their properties. In particular, such properties may be quantified over, may be values of variables, and as such may be arguments of relations. This extension seems to be needed to implement the account of unlike category coordination sketched in Yatabe 2004. This improvement is outlined in Section 4, where I also motivate it by briefly arguing that there is currently no viable HPSG alternative to Yatabe's (2004) analysis.

These three improvements are orthogonal in the sense that any of them may be adopted, without the need to adopt any of the others. Accordingly, each of the following three sections may be read independently of the others.

2 Non-Exhaustive Rooted Models

Since King 1999, HPSG models are assumed to be exhaustive (see Richter 2004, 2007; cf. Pollard 1999), i.e., contain all possible kinds of structures licensed by the grammar. For example, a single HPSG model of English will contain structures for all possible English utterances and words, as well as many partial structures satisfying the grammar (e.g., various *local* or *synsem* objects). This corresponds to the intuition that grammars describe whole languages, so each model should represent the whole language. However, there is another valid intuition, which is predominant outside of HPSG: that grammars describe possible utterances. This latter intuition leads to much smaller models: each model corresponds to a single utterance and only the collection of all models corresponds to the whole language.

By way of analogy, consider the artificial toy problem of describing all configurations of black and white objects such that each black object is related to at least one white object and vice versa (cf. Przepiórkowski 2021: § 4). The following first order formulae are a reasonable theory of such configurations:

(1) $\forall x.\, \mathit{black}(x) \leftrightarrow \neg \mathit{white}(x)$

(2) $\forall x \forall y.\, \mathit{bw}(x, y) \rightarrow \mathit{black}(x) \wedge \mathit{white}(y)$

(3) $\forall x.\, \mathit{black}(x) \rightarrow \exists y.\, \mathit{white}(y) \wedge \mathit{bw}(x, y)$

(4) $\forall x.\, \mathit{white}(x) \rightarrow \exists y.\, \mathit{black}(y) \wedge \mathit{bw}(y, x)$

Together they are saying that everything is either *black* or *white* (see (1)) and that there is a relation, *bw*, which holds between *black* things and *white* things (see (2)) such that every *black* thing is in this relation with some (at least one) *white* thing (see (3)) and every *white* thing is related to some (at least one) *black* thing (see (4)). There are models of this theory of any cardinality apart from 1 (including transfinite cardinalities): the empty model satisfies (1)–(4) and so does, e.g., any model which contains exactly one *white* thing and arbitrarily many (but at least one) *black* things appropriately related to it. Now imagine that, as in HPSG, models were required to be exhaustive, i.e., each model would have to contain all possible configurations of white and black objects. It is not clear what such models would contribute

to our understanding of the described black and white configurations above the simpler non-exhaustive models, but it is clear that they would be dubious from the point of view of the standard (ZFC) set theory: such models would be too large to be sets.[1]

Also in the case of HPSG, exhaustive models lead to some serious problems, discussed in Richter 2007. One, dubbed *twin structures*, is that some parts of the model might simultaneously belong to two different utterances, which does not correspond to any empirical facts. Another, called *stranded structures*, is that models may contain structures smaller than utterances (e.g., certain structures rooted in *local* objects), including structures (called *stranded monster structures* in Richter 2007) which may never be parts of any utterances and which are intuitively clearly ill-formed. Richter 2007 retains the idea of exhaustive models and deals with these problems by imposing restrictions on HPSG signatures, to the effect that all sorts (including such formerly atomic sorts as *nom* or *sg*) are specified for the attribute EMBEDDED, whose value is an unembedded sign (*u_sign*):

(5) *top* EMBEDDED *u_sign*
  sign ...
    e_sign ...
    u_sign ...
  ...

Moreover, there is just one *u_sign* object – an unembedded sign – in each configuration of objects (see (6)) and all objects in a configuration are components of this unembedded sign (see (7)).

(6) UNIQUE U-SIGN CONDITION:
$\forall\boxed{1}\forall\boxed{2}((\boxed{1} \sim u\_sign \wedge \boxed{2} \sim u\_sign) \rightarrow \boxed{1} \approx \boxed{2})$

(7) U-SIGN COMPONENT CONDITION:
$\forall\boxed{1}(\boxed{1} \sim top \rightarrow \exists\boxed{2}(\boxed{2} \sim u\_sign \wedge \texttt{component}(\boxed{1}, \boxed{2})))$

The combined effect of (5)–(7) is that each configuration of objects in an exhaustive model contains exactly one unembedded sign that all these objects are components of (i.e., are reachable from); this unembedded sign acts as the root of an utterance.

This solves the two problems identified in Richter 2007. There are no *twin structures*, as each object is a component of just a single *u_sign*. There are also no *stranded structures*, on the assumption that *u_sign* is appropriately constrained to the effect that its SLASH value is empty, its VALENCE lists are empty, etc. However, this solution comes at a considerable cost: not only are all structures massively cyclic (each object has the attribute EMBEDDED whose value is the utterance to which this object belongs), but there is also the conceptual problem of, say, the value of CASE containing the whole utterance

[1] In brief, they would contain configurations of arbitrarily large cardinality, so they themselves would not have any cardinality (as there is no maximal cardinality).

(given that, e.g., the sort *nom* is specified for the attribute EMBEDDED). In any case, this solution leads to very different structures than what HPSG linguists are used to.

Richter 2007: 102 claims that the problem of *stranded monster structures* arises because "[t]he grammars in the HPSG literature are not precise enough for their models to match the intentions of linguists". (This justifies the solution alluded to above, consisting in the modifications of the grammar rather than the model theory.) However, it would be unrealistic to expect of linguists to be aware of – and deal with – such technical model-theoretic problems. So a better diagnosis of the problems mentioned above is that they arise because the HPSG model theory does not sufficiently meet the needs of linguists, who only care about utterances and their components, and do not intend their grammars to say anything about, for example, arbitrary objects of sort *local* outside of utterances.

The crucial observation is that all the problems identified in Richter 2007 disappear when a leaner approach to modelling is adopted, upon which each model corresponds to a single utterance, as commonly assumed elsewhere. Specifically, I propose that HPSG models be *rooted* (*point generated*) in the sense of modal logic:[2] one object of the universe is singled out and it serves as the root of the model. This object may be referred to directly in HPSG descriptions via a special symbol, r.

In order to make sure that the distinguished object is really the root of the whole model, the following constraint must be present in each HPSG grammar, where `component` is defined in the standard way (e.g., Sailer 2003: 115–116):

(8) $\forall \boxed{1}$ `component`$(\boxed{1}, r)$

This states that each object in the model is reachable from the distinguished object via some sequence of attributes.

One immediate advantage of this approach is that it makes it easy to state constraints on utterances. For example, the requirement that utterances have empty SLASH may be stated directly as in (9) (assuming that empty sets are modelled via objects of sort *eset*; Richter 2004: 281), without the need for technical boolean attributes such as ROOT (e.g., in Ginzburg & Sag 2000).

(9) r NONLOCAL SLASH $\sim$ *eset*

Full technical details are given in Appendix A. Here let me only point out that this simple view of HPSG models as rooted models solves the problems addressed in Richter 2007. There are no *twin structures*, as each model corresponds to a single utterance, and there are no *stranded structures* (*monster* or not), as each structure is a part of an utterance. Unlike the proposal in Richter 2007, this solution does not require extensions of signatures and does not result in rather different models than what HPSG linguists are used to, ones that have the cyclicity-inducing EMBEDDED attribute defined

[2] See, e.g., Blackburn et al. 2010: 56, 107; cf. *singly generated* models in Pollard 1999.

on every sort.

On a more conceptual note, rooted non-exhaustive models proposed here are also compatible with all views on the nature of the objects residing in such models: they may be understood as abstract feature structures (Pollard & Sag 1994) or other mathematical idealisations of types of utterances (Pollard 1999), but they may also be understood as utterance tokens, as in King 1999. On the latter view, there is a tension between the idea that model objects are specific linguistic tokens and the idea that models are exhaustive, i.e., contain all configurations that the grammar predicts. Clearly, any realistic grammar predicts the grammaticality of certain utterance types that have never been – and never will be – actually uttered, i.e., utterance types for which there are no actual tokens. This forces King (1999) to assume "non-actual tokens", a concept that may be considered "contradictory and nonsensical" (Richter 2004: 119, citing Carl Pollard, p.c.). Giving up exhaustivity makes it possible to adopt King's (1999) view on the nature of model objects as utterance tokens.

Let me finally point out the affinity of the proposed solution with Pollard's (1999: §6) *singly generated models*. In both cases, one object in a model is distinguished as root, but Pollard (1999) does not require that this object be the root of an utterance (nor is it possible to refer to this object directly in the grammar). This makes the approach of Pollard 1999 – but not the approach proposed here – susceptible to some of the problems discussed in Richter 2007.

3 Höhle's Problem

Höhle's Problem is similar to the problems discussed in the previous section in the sense that it is concerned with the fact that there are configurations in models which are not expected by linguists, but it differs in that these configurations are not exactly wrong: rather, they are spurious and there are many, many more of them than desired. Let us illustrate the problem with the following sentence:

(10) She says she loves you.

From the linguistic point of view, there just two different analyses of this sentence: one in which INDEX values of the two pronouns *she* are token-identical, and one in which they are not.[3] That is, in the model of English, there are two configurations corresponding to (10), fragmentarily represented in Figure 1, which differ only in whether $\boxed{1} = \boxed{2}$ or $\boxed{1} \neq \boxed{2}$.

Let us concentrate on one of these, say, on the one in which the two pronouns *she* are not coindexed. The problem – Höhle's Problem – is that

[3]Typial HPSG representations of this sentence will also differ in values of GEND(er) and NUM(ber) within the INDEX value of the pronoun *you*, but – as discussed in Przepiórkowski 2021: § 3.3.2 – it is not clear whether having such ambiguities is desirable.

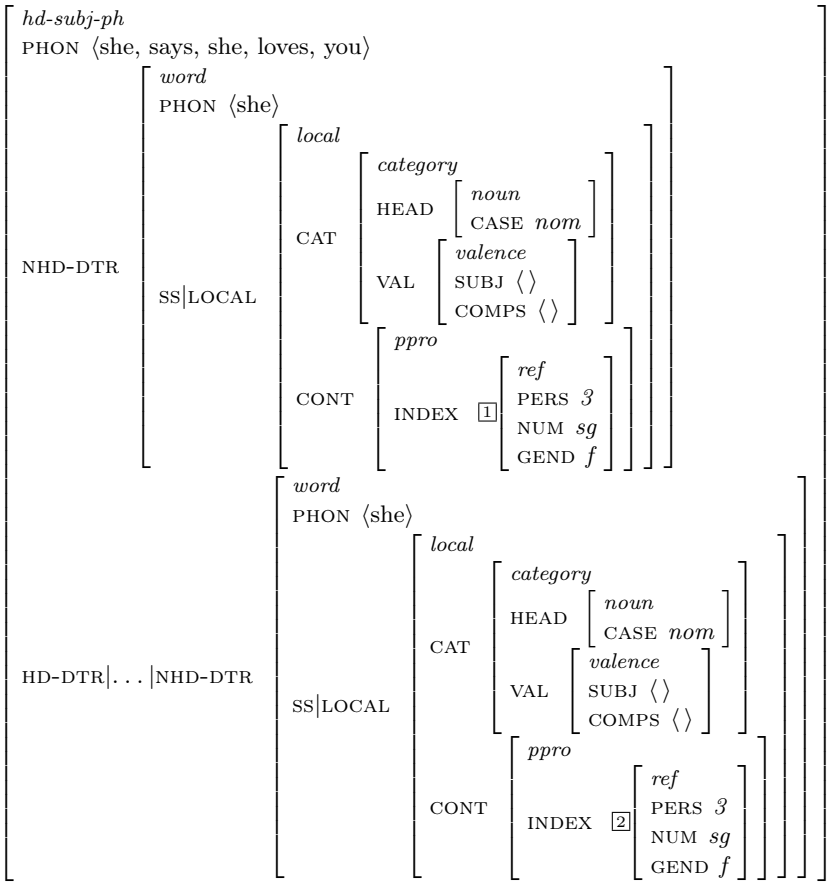

Figure 1: Fragments of an AVM representation of (10)

there are still many different configurations in the model corresponding to Figure 1 with $\boxed{1} \neq \boxed{2}$, which differ in ways that linguists do not suspect and certainly do not care about. For example, even if the two INDEX values are different objects, the values of any of the attributes within INDEX may be token-identical or not. One possibility is schematically shown in Figure 2, where the two INDEX values corresponding to the two pronouns *she* are different model objects of sort *ref* (objects 1 and 5), and GEND values are also different objects of sort *f* (objects 4 and 6), but the values of the two attributes PERS are the same object 2 of sort *3*, and the values of the two attributes NUM are the same object 3 of sort *sg*. It is easy to see that there are $2^3 = 8$ different configurations corresponding to two non-token-identical INDEX values of the two pronouns *she* in (10). But of course this is just the tip of the iceberg. In model configurations corresponding to the schematic representation in Figure 1, the two CAT values may be identical or not; if they are not, HEAD values might be the same object or not; and if they are not, CASE values may be token-identical or not. Similarly for VAL

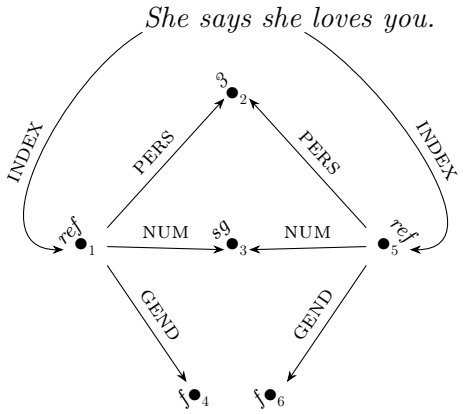

Figure 2: One of the eight different configurations of non-token-identical INDEX values of the two pronouns *she* in (10)

values, HEAD values of the two verbs, etc. Even if we ignore the multiple occurrences of empty lists in the structure, there are thousands of different model configurations corresponding to the sentence in (10) with the two feminine pronouns non-coindexed (and only eight times fewer when they are coindexed). As observed in Przepiórkowski 2021, adding to the equation the problem of which empty list values of various attributes (e.g., the two attributes SUBJ and the two attributes COMPS in Figure 1, among many others) are the same *elist* object and which are different *elist* objects, results in literally billions of different configurations where the linguist would expect just one.[4]

There are partial solutions of Höhle's Problem in the literature. Richter (2007: 102) proposes the Unique Empty List Condition in (11), which makes sure that all empty lists within an utterance are the same *elist* object.

(11) $\forall\boxed{1}\forall\boxed{2}\,((\boxed{1}\sim \mathit{elist} \wedge \boxed{2}\sim \mathit{elist}) \rightarrow \boxed{1}\approx\boxed{2})$

A comprehensive principle, which says roughly that structures that look the same are token-identical, is Sailer's (2003: 116) General Identity Principle (GIP) in (12), with the definition of the relation `are-copies` given in (13).

(12) $\forall\boxed{1}\forall\boxed{2}\,(\texttt{are-copies}(\boxed{1},\boxed{2}) \rightarrow \boxed{1}\approx\boxed{2})$

(13) $\forall\boxed{1}\forall\boxed{2}\ \texttt{are-copies}(\boxed{1},\boxed{2}) \leftrightarrow$

$$\left(\bigvee_{\sigma\in\mathcal{S}} (\boxed{1}\sim\sigma \wedge \boxed{2}\sim\sigma) \wedge \bigwedge_{\alpha\in\mathcal{A}} (\boxed{1}\alpha \approx \boxed{1}\alpha \rightarrow \texttt{are-copies}(\boxed{1}\alpha, \boxed{2}\alpha))\right)$$

[4] "Each word introduces three lists (values of PHON, VAL|SUBJ, and VAL|COMPS), and there are five words in this sentence, so there are 15 *elist* objects stemming from words alone. The number of different ways to partition a set of n elements into equivalence classes is given by Bell numbers B_n, and B_{15} = 1,382,958,545 (see https://oeis.org/A000110/list). This should be multiplied by the eight configurations of the two [INDEX values], etc." (Przepiórkowski 2021: fn. 44).

In (13), $\mathcal{S}$ stands for the set of species (i.e., maximally specific sorts) and $\mathcal{A}$ – for the set of attributes. What (13) is saying is that two objects $\boxed{1}$ and $\boxed{2}$ are copies if and only if they are of the same species σ and, recursively, for any attribute α defined for that species, the values of α for $\boxed{1}$ and for $\boxed{2}$ are copies.

As stated here, GIP is too strong: it requires that INDEX values of the two pronouns *she* in (10) must be token-identical, i.e., it invalidates the analysis of (10) on which the two feminine pronouns are not coindexed. More generally, this GIP is incompatible with the standard HPSG binding theory, which requires that some same-looking INDEX values are not token-identical. For this reason, GIP is formulated in Sailer 2003 in such a way that it only applies to certain semantic representations, in a way that is compatible with the standard binding theory and preserves the ambiguity of (10). But this means that the problem of spurious ambiguities remains. What is required to solve Höhle's Problem is a way to constrain the scope of GIP more selectively, for example, a way to say that it must apply to all same-looking structures with the exception of INDEX values of sort *ref*.[5]

The rest of this section describes a relatively simple solution, one that is much more comprehensive than the Unique Empty List Condition in (11) or the General Identity Principle in (12) constrained to certain semantic representations, but still leaves the theoretical possibility of spurious ambiguities occurring in some very special cases. A fully general but more complex solution is presented in Appendix B.

The key observation in the simpler solution is that the definition of `are-copies` in (13) does not determine whether same-looking *cyclic* structures stand in this relation or not. I will demonstrate the correctness of this observation below. But if such same-looking structures are not in the `are-copies` relation, then GIP in (12) does not force them to be token-identical. This means that one way to make GIP fully general but still allow, say, for INDEX values of sort *ref* to escape GIP, is to make such INDEX values cyclic. This can be achieved by adding one more attribute to *ref*, let us call it INT for *intensional*,[6] as in the signature fragment in (14), and by making sure – via the constraint in (15) – that the value of INT is the object on which this attribute occurs, in effect creating a small cycle.

(14) *ref* GEND *gender*
NUM *number*
PERS *person*
INT *ref*

(15) Universal INTENSIONALITY PRINCIPLE:
$\forall\boxed{1}\forall\boxed{2}\,(\,\boxed{1}\text{INT} \approx \boxed{2} \rightarrow \boxed{1} \approx \boxed{2}\,)$

[5] Other, less broadly accepted analyses which rely on some same-looking structures not being token identical, are Höhle's (1999: §2.4) architecture for phonology and Meurers's (1998: 326, fn. 42) approach to structural case assignment.

[6] Thanks to Frank Richter (p.c.) for suggesting this name.

With this modification in hand, let us see whether two INDEX values in (16) for the two occurrences of *she* – which intuitively look the same – are in the `are-copies` relation according to its definition (13).

(16) $\boxed{1}\begin{bmatrix} \textit{ref} \\ \text{PERS } \textit{3} \\ \text{NUM } \textit{sg} \\ \text{GEND } \textit{f} \\ \text{INT } \boxed{1} \end{bmatrix} \overset{?}{\approx} \boxed{2}\begin{bmatrix} \textit{ref} \\ \text{PERS } \textit{3} \\ \text{NUM } \textit{sg} \\ \text{GEND } \textit{f} \\ \text{INT } \boxed{2} \end{bmatrix}$

According to this definition, 1 and 2 in (16) are in the `are-copies` relation iff 1) they are of same species (*yes* – both are of species *ref*), 2) the values of PERS are copies (*yes* – they are of the same species *3* and have no attributes), 3) the values of NUM are copies (*yes*), 4) the values of GEND are copies (*yes*), and 5) the values of INT are copies. That is, 1 and 2 *qua* values of INDEX are copies if and only if 1 and 2 *qua* values of INT are copies. In other words, the definition of `are-copies` does not determine whether 1 and 2 are copies. Since they do not have to be in the `are-copies` relation, they are not forced by GIP to be the same objects. That is, they are genuinely exempt from GIP, even though they look the same.

In summary, the proposed simpler solution to Höhle's Problem consists in 1) adopting Sailer's (2003: 116) General Identity Principle but without restricting its scope to semantic representations or any other specific configurations and in 2) making structures that should be exempt from GIP cyclic.

This is a much more general solution than the partial solutions mentioned above. It subsumes Richter's (2007: 102) Unique Empty List Condition in (11), as it makes not only empty lists but all attribute-less species unique, so that in any utterance there is only one *elist* object, at most one *nom* object, at most one *sg*, etc. It also subsumes GIP as understood in Sailer 2003, since it is applied there to configurations which are not cyclic. However, this solution is not completely general, as it makes *all* cyclic structures exempt from GIP. So, for example, in a grammar of English in which a determiner and a noun mutually select each other, there will typically be a cycle in each nominal phrase containing a determiner. In such a case, when the structures of two NPs look the same (e.g., *the guy* in the sentence *The guy's mother loves the guy's father*), some spurious ambiguities will occur despite GIP. A more elaborate solution that is fully general and does not rely on cyclicity is presented Appendix B.

4 Coordination of Unlikes: Second-Order HPSG

In order to handle examples such as (17) (from Bayer 1996: 585, fn. 7, (ii.c–d)), Yatabe (2004: 343) assumes a lexical entry for *emphasized* schematically represented in (18), with the category of the object specified disjunctively as an NP (nominal phrase; see *noun*) or a CP (complementiser phrase; *comp*).

(17) a. We emphasized [[Mr. Colson's many qualifications]$_{\text{NP}}$ and [that he had worked at the White House]$_{\text{CP}}$].

b. We emphasized [[that Mr. Colson had worked at the White House]$_{\text{CP}}$ and [his many other qualifications]$_{\text{NP}}$].

(18)
$$\left[\begin{array}{l}\textsc{phon}\ \langle\text{emphasized}\rangle \\ \ldots\ \textsc{valence}\ \left[\begin{array}{l}\textsc{subj}\ \langle\left[\ \ldots\ \textsc{head}\ c(\left[\begin{array}{l}\textit{noun}\\ \textsc{case}\ \textit{nom}\end{array}\right])\ \right]\rangle \\ \textsc{comps}\ \langle[\ \ldots\ \textsc{head}\ c(\textit{noun}\vee\textit{comp})\ ]\rangle\end{array}\right]\end{array}\right]$$

The key idea is the use of the distributive functor, c, defined in (19) (Yatabe 2004: 343, (12)):

(19) $\boxed{1} : c(\alpha) \equiv \boxed{1} : \alpha \vee (\boxed{1} : [\ \textsc{args}\ \langle\boxed{a_1},\ldots,\boxed{a_n}\rangle\ ] \wedge \boxed{a_1} : \alpha \wedge \ldots \wedge \boxed{a_n} : \alpha)$

Here α is a description, such as $\left[\begin{array}{l}\textit{noun}\\ \textsc{case}\ \textit{nom}\end{array}\right]$ or $\textit{noun} \vee \textit{comp}$ in (18), and an object $\boxed{1}$ satisfies $c(\alpha)$ – written as $\boxed{1} : c(\alpha)$ – iff it either satisfies the description α directly (see the first disjunct in (19)), or if it is the HEAD value of a coordinate structure with conjuncts having HEAD values $\boxed{a_1}$, …, $\boxed{a_n}$ (see the second disjunct); in the latter case, each of $\boxed{a_1}$, …, $\boxed{a_n}$ must satisfy α independently.

The intention of (19) is clear, but it is far from clear how to formally encode it. That is, for each particular description α it is easy to define a unary relation corresponding to $c(\alpha)$ in (19). What is far from clear is how to define c in its generality (i.e., in a way simulating (19)), as a binary relation between objects and arbitrary descriptions α. The problem is that, in RSRL (Relational Speciate Re-entrant Language; Richter 2004), the language for formalising HPSG grammars, arguments of relations are objects, not descriptions.

I argue that this kind of analysis of unlike category coordination (UCC) is on the right track – to the extent that justifies making RSRL a second-order language, in which not only objects but also their properties may be quantified over.[7] While linearisation-based approaches to UCC were popular in HPSG in 2000s (e.g., Crysmann 2003, Beavers & Sag 2004, Chaves 2006, 2008), it is clear now that at least some cases of UCC must be analysed as direct coordination of smaller constituents, rather than as coordination of larger verbal constituents and subsequent ellipsis (see, e.g., Levine 2011: § 2.3, Dalrymple 2017, Abeillé & Chaves 2021: § 6, and Patejuk & Przepiórkowski 2021). Conceding this point, Chaves 2013 proposes to save the law of the coordination of likes (as it is sometimes called after Williams 1981) by reanalysing categories as constellations of some morphosyntactic features and moving troublesome distributive restrictions, such as those encoded in (18), to semantics. Unfortunately, this approach is untenable, given that CASE is one of the remaining categorial features in Chaves 2013 and that instances

[7] Second-order systems usually have higher computational complexity than their first-order equivalents, but given that already first-order RSRL is undecidable (Kepser 2004), second-order RSRL is in the same class as standard RSRL.

of unlike case coordination are well known (and have also been discussed within HPSG; see Przepiórkowski 1999: § 5.3.1 and Levy 2001: § 4). Hence, Yatabe's (2004) is the most convincing approach to UCC currently on the HPSG market and, given that the distributive functor c is also explicitly invoked in recent work (Yatabe & Tam 2021: 74), there is an increasing need to make it formalisable.

This calls for extending the syntax and semantics of RSRL to handle second-order quantification. The modifications of the standard RSRL definitions are relatively straightforward:[8]

- signatures do not only specify arities of relation symbols, but also types of their arguments (each either e or et);
- interpretations of relation symbols are trivially modified so that they satisfy such signatures (i.e., they are sets of tuples whose each element is an object or a set of objects, depending on the type specified in the signature);
- the set of variables, $\mathcal{VAR}$, is the disjoint sum of $\mathcal{VAR}_e$ (first-order variables) and $\mathcal{VAR}_{et}$ (second-order variables);
- variable assignments assign objects to elements of $\mathcal{VAR}_e$ and they assign sets of objects to elements of $\mathcal{VAR}_{et}$; the interpretation of quantifiers is extended to second-order variables correspondingly;
- apart from the usual first-order terms $\mathcal{T}_e^\Sigma$ (for the signature Σ), there are also second-order terms, $\mathcal{T}_{et}^\Sigma$, specified recursively simultaneously with the set of formulae, $\mathcal{D}^\Sigma$, as the disjoint sum of second-order variables ($\mathcal{VAR}_{et}$) and all formulae ($\mathcal{D}^\Sigma$);
- two clauses of the definition of formulae (Richter 2004: 165) are further modified so that:
  - the variables which are arguments of relation symbols are of the right type e or et,
  - $\tau_1 \approx \tau_2$ is a formula if both terms are of the same type (i.e., both are e or both are et);
- importantly, a new kind of formula is added: $\tau_1(\tau_2)$, where $\tau_1 \in \mathcal{T}_{et}^\Sigma$ and $\tau_2 \in \mathcal{T}_e^\Sigma$; this formula says that the description τ_1 holds of the object τ_2;
- more precisely, the interpretation of $\tau_1(\tau_2)$ is the set of all these objects of the universe U on which the interpretation of τ_2 belongs to the interpretation of τ_1; more formally: $D_{\mathsf{I}}^{ass}(\tau_1(\tau_2)) = \{u \in \mathsf{U} : T_{\mathsf{I}}^{ass}(\tau_2)(u) \in D_{\mathsf{I}}^{ass}(\tau_1)\}$.

Note that, apart from the extended interpretations of relation symbols, models are not affected by these changes: they are still collections of objects

[8] See Richter 2004: § 3.1.1 for the standard definitions and meanings of particular symbols. I simplify throughout by ignoring chains.

of particular species related via particular attributes.

Given these extensions, the lexical entry in (18) may be represented as in (20), with the definition of c in (19) formalised via the relation $\mathtt{c}$ defined in (21).

(20)
$$\left[\begin{array}{l}\textsc{phon}\ \langle\text{emphasized}\rangle \\ \ldots\ \textsc{valence}\ \left[\begin{array}{l}\textsc{subj}\ \langle[\ \ldots\ \textsc{head}\ \boxed{1}\ ]\rangle \\ \textsc{comps}\ \langle[\ \ldots\ \textsc{head}\ \boxed{2}\ ]\rangle\end{array}\right]\end{array}\right] \wedge \alpha_1 \approx (: \sim \mathit{noun} \wedge : \textsc{case} \sim \mathit{nom}) \\ \wedge\ \alpha_2 \approx (: \sim \mathit{noun} \vee : \sim \mathit{comp}) \\ \wedge\ \mathtt{c}(\boxed{1}, \alpha_1) \wedge \mathtt{c}(\boxed{2}, \alpha_2)$$

(21)
$$\forall \boxed{1}_e \forall \alpha_{et}\ (\ \mathtt{c}(\boxed{1}, \alpha) \leftrightarrow \alpha(\boxed{1}) \vee \exists \boxed{a_1} \ldots \exists \boxed{a_n}\ (\ \boxed{1}[\ \textsc{args}\ \langle \boxed{a_1}, \ldots, \boxed{a_n}\rangle\ ] \wedge \mathtt{c}(\boxed{a_1}, \alpha) \wedge \ldots \wedge \mathtt{c}(\boxed{a_n}, \alpha)\ )\ )$$

The definition of relation $\mathtt{c}$ in (21) differs from Yatabe's (2004) definition of c in (19) in being fully recursive, i.e., in taking into account nested (embedded) coordination, as in *Scooby-Doo or Tom and Jerry*.

As already pointed out above, this second-order extension of RSRL seems to be necessary to formalise Yatabe's (2004) analysis in its generality. What may be considered an advantage of this formalisation is that it also encodes the standard LFG approach to coordination, on which certain properties are distributive so that, when they are applied to a coordinate structure, they independently distribute to all conjuncts (see, e.g., Dalrymple & Kaplan 2000, Przepiórkowski & Patejuk 2012, and especially Przepiórkowski & Patejuk 2021). That it, the second-order extension proposed here makes it possible to formally define the notion of distributivity in coordination which is assumed in Lexical Functional Grammar as a primitive mechanism of that theory.

It must be noted, however, that an extensionally equivalent analysis – i.e., an analysis that results in exactly the same configurations in models – is possible that does not require such a second-order extension: instead of defining the second-order relation $\mathtt{c}$ whose second argument is an arbitrary description, it is possible to define a different first-order relation for each such description. For example, the lexical entry in (20) may be replaced with the lexical entry in (22), with relations `noun_and_nom` and `noun_or_comp` defined as in (23)–(24):

(22)
$$\left[\begin{array}{l}\textsc{phon}\ \langle\text{emphasized}\rangle \\ \ldots\ \textsc{valence}\ \left[\begin{array}{l}\textsc{subj}\ \langle[\ \ldots\ \textsc{head}\ \boxed{1}\ ]\rangle \\ \textsc{comps}\ \langle[\ \ldots\ \textsc{head}\ \boxed{2}\ ]\rangle\end{array}\right]\end{array}\right] \wedge \mathtt{noun\_and\_nom}(\boxed{1}) \\ \wedge\ \mathtt{noun\_or\_comp}(\boxed{2})$$

(23)
$$\forall \boxed{1}\ (\ \mathtt{noun\_and\_nom}(\boxed{1}) \leftrightarrow (\ (\boxed{1} \sim \mathit{noun} \wedge \boxed{1}\textsc{case} \sim \mathit{nom}) \vee \exists \boxed{a_1} \ldots \exists \boxed{a_n}\ (\ \boxed{1}[\ \textsc{args}\ \langle \boxed{a_1}, \ldots, \boxed{a_n}\rangle\ ] \wedge \mathtt{noun\_and\_nom}(\boxed{a_1}) \wedge \ldots \wedge \mathtt{noun\_and\_nom}(\boxed{a_n})\ )\ )\ )$$

(24)
$$\forall \boxed{1}\ (\ \mathtt{noun\_or\_comp}(\boxed{1}) \leftrightarrow (\ (\boxed{1} \sim \mathit{noun} \vee \boxed{1} \sim \mathit{comp}) \vee \exists \boxed{a_1} \ldots \exists \boxed{a_n}\ (\ \boxed{1}[\ \textsc{args}\ \langle \boxed{a_1}, \ldots, \boxed{a_n}\rangle\ ] \wedge \mathtt{noun\_or\_comp}(\boxed{a_1}) \wedge \ldots \wedge \mathtt{noun\_or\_comp}(\boxed{a_n})\ )\ )\ )$$

As different predicates impose different selectional restrictions and allow for different combinations of categories, many relations analogous to (23)–(24) would have to be defined in the grammar, all encoding essentially the same

mechanism of distribution of selectional restrictions to all conjuncts in a coordinate structure. For this reason, an analysis in terms of a single general relation encoding such distributivity should be preferred, even if it calls for a second-order extension of RSRL.

5 Conclusion

While the extent to which the model theory of HPSG is developed is unparalleled, and – with the notable exception of Søgaard & Lange 2009 – there is practically no work on the formal foundations of HPSG after Richter 2007, it would be a mistake to assume that all problems are solved and all reasonable analyses may be formalised. The improvements proposed in this paper range from fundamental and conceptual (making models rooted and non-exhaustive, extending the underlying language to second-order) to purely technical (solving the long-standing Höhle's Problem). I hope that this paper will help rekindle some interest in the formal foundations of HPSG.

Appendices

A Non-Exhaustive Rooted Models – Technicalities

Here are the technical modifications to RSRL, as defined in Richter 2004: § 3.1.1, which are needed to implement the idea of rooted non-exhaustive models presented in Section 2. All definitions are simplified by ignoring complications related to chains.

I assume the standard notion of *signature* (Richter 2004: 156):

Definition 1 (signature) *Σ is a signature iff*
Σ is a septuple $\langle S, \sqsubseteq, S_{max}, A, F, R, Ar\rangle$,
$\langle S, \sqsubseteq\rangle$ is a partial order,
$S_{max} = \{\sigma \in S \mid$ for each $\sigma' \in S$, if $\sigma' \sqsubseteq \sigma$ then $\sigma = \sigma'\}$,
A is a set,
F is a partial function from $S \times A$ to S,
for each $\sigma_1 \in S$, for each $\sigma_2 \in S$, for each $\phi \in A$,
if $F(\sigma_1, \phi)$ is defined and $\sigma_2 \sqsubseteq \sigma_1$
then $F(\sigma_2, \phi)$ is defined and $F(\sigma_2, \phi) \sqsubseteq F(\sigma_1, \phi)$,
R is a finite set, and
Ar is a total function from R to the positive integers.

On the other hand, I extend the notion of *terms* (Richter 2004: 162) by adding a special symbol, r, used to refer to the distinguished object in the universe of a interpretation (which will be defined below, in Definition 5):[9]

[9] In this and the following definitions, my extensions are underlined.

Definition 2 (terms) *For each signature* $\Sigma = \langle S, \sqsubseteq, S_{max}, A, F, R, Ar\rangle$, *the set of terms* T^Σ *is the smallest set such that*
$\underline{r \in T^\Sigma}$,
$: \in T^\Sigma$,
for each $x \in V$, $x \in T^\Sigma$,
for each $\phi \in A$ *and each* $\tau \in T^\Sigma$, $\tau\phi \in T^\Sigma$.

The definition of *formulæ* is standard (Richter 2004: 165):

Definition 3 (formulæ) *For each signature* $\Sigma = \langle S, \sqsubseteq, S_{max}, A, F, R, Ar\rangle$, *the set of formulæ* D^Σ *is the smallest set such that*
for each $\sigma \in S$, *for each* $\tau \in T^\Sigma$, $\tau \sim \sigma \in D^\Sigma$,
for each $\tau_1, \tau_2 \in T^\Sigma$, $\tau_1 \approx \tau_2 \in D^\Sigma$,
for each $\rho \in R$, *for each* $x_1, \ldots, x_{Ar(\rho)} \in V$, $\rho(x_1, \ldots, x_{Ar(\rho)}) \in D^\Sigma$,
for each $x \in V$, *for each* $\delta \in D^\Sigma$, $\exists x\delta \in D^\Sigma$, *(analogous for* $\forall$*)*
for each $\delta \in D^\Sigma$, $\neg\delta \in D^\Sigma$,
for each $\delta_1, \delta_2 \in D^\Sigma$, *and* $(\delta_1 \wedge \delta_2) \in D^\Sigma$. *(analogous for* $\vee, \rightarrow, \leftrightarrow$*)*

Additionally, the standard definition of *free variables* (Richter 2004: 166–167), FV, is trivially extended so that the term r is variable-free: $FV(r) = \{\}$.

Also the definition of *descriptions* is standard (Richter 2004: 173):

Definition 4 (descriptions) *For each signature* Σ, *the set of descriptions* $D_0^\Sigma = \left\{\delta \in D^\Sigma | FV(\delta) = \{\}\right\}$.

The definition of *interpretation* is extended from a quadruple $\langle \mathsf{U}, \mathsf{S}, \mathsf{A}, \mathsf{R}\rangle$ (Richter 2004: 157–158) to a quintuple $\langle \mathsf{U}, \mathsf{r}, \mathsf{S}, \mathsf{A}, \mathsf{R}\rangle$, where U, S, A, and R are defined in the standard way (i.e., as the universe, assignment of species to objects, interpretation of attributes, and interpretation of relation symbols, respectively), and $\mathsf{r} \in \mathsf{U}$ is the distinguished object:

Definition 5 (interpretation) *For each signature*
$\Sigma = \langle S, \sqsubseteq, S_{max}, A, F, R, Ar\rangle$, $\mathsf{I} = \langle \mathsf{U}, \underline{\mathsf{r}}, \mathsf{S}, \mathsf{A}, \mathsf{R}\rangle$ *is an* Σ *interpretation iff*
U *is a set,*
$\underline{\mathsf{r} \in \mathsf{U}}$,
S *is a total function from* U *to* S_{max},
A *is a total function from* A *to the set of partial functions from* U *to* U,
for each $\phi \in A$ *and each* $u \in \mathsf{U}$
  if $\mathsf{A}(\phi)(u)$ *is defined*
  then $F(\mathsf{S}(u), \phi)$ *is defined, and* $\mathsf{S}(\mathsf{A}(\phi)(u)) \sqsubseteq F(\mathsf{S}(u), \phi)$, *and*
for each $\phi \in A$ *and each* $u \in \mathsf{U}$,
  if $F(\mathsf{S}(u), \phi)$ *is defined then* $\mathsf{A}(\phi)(u)$ *is defined,*
R *is a total function from* R *to the power set of* $\bigcup_{n\in\mathbb{N}} \mathsf{U}^n$, *and*
  for each $\rho \in R$, $\mathsf{R}(\rho) \subseteq \mathsf{U}^{Ar(\rho)}$.

The definition of *variable assignments* (Richter 2004: 161–162) is standard, while the definition of *term interpretation* (Richter 2004: 162–163) is extended so that the interpretation of the term r is the distinguished object r:

Definition 6 (term interpretation) *For each signature* $\Sigma = \langle S, \sqsubseteq, S_{max}, A, F, R, Ar\rangle$*, for each* Σ *interpretation* $\mathsf{I} = \langle \mathsf{U}, \mathsf{r}, \mathsf{S}, \mathsf{A}, \mathsf{R}\rangle$*, for each* $g \in \mathsf{G_I}$*, the term interpretation* T_I^g *is the total function from* T^Σ *to the set of partial functions from* U *to* U *such that for each* $u \in \mathsf{U}$*,*

$$\frac{\mathsf{T}_\mathsf{I}^g(r)(u) \text{ is defined and } \mathsf{T}_\mathsf{I}^g(r)(u) = \mathsf{r},}{\mathsf{T}_\mathsf{I}^g(:)(u) \text{ is defined and } \mathsf{T}_\mathsf{I}^g(:)(u) = u,}$$

for each $x \in V$*,* $\mathsf{T}_\mathsf{I}^g(x)(u)$ *is defined and* $\mathsf{T}_\mathsf{I}^g(x)(u) = g(x)$*,*
for each $\tau \in T^\Sigma$*, for each* $\phi \in A$*,*
$\mathsf{T}_\mathsf{I}^g(\tau\phi)(u)$ *is defined iff* $\mathsf{T}_\mathsf{I}^g(\tau)(u)$ *is defined and* $\mathsf{A}(\phi)(\mathsf{T}_\mathsf{I}^g(\tau)(u))$ *is defined, and*
if $\mathsf{T}_\mathsf{I}^g(\tau\phi)(u)$ *is defined then* $\mathsf{T}_\mathsf{I}^g(\tau\phi)(u) = \mathsf{A}(\phi)(\mathsf{T}_\mathsf{I}^g(\tau)(u))$*.*

The definition of *formula denotation* (Richter 2004: 168–169) can be simplified: given that the whole universe in any interpretation corresponds to a single utterance, quantification may now by defined in the same way as in first-order logic, as quantification over the whole universe, rather than as quantification over components. The practical effect of this modification is the same as in the setup of Richter 2007, where quantification evaluated at any object scopes over the whole utterance to which this object belongs.

Definition 7 (formula denotation) *For each signature* $\Sigma = \langle S, \sqsubseteq, S_{max}, A, F, R, Ar\rangle$*, for each* Σ *interpretation* $\mathsf{I} = \langle \mathsf{U}, \mathsf{r}, \mathsf{S}, \mathsf{A}, \mathsf{R}\rangle$*, for each* $g \in \mathsf{G_I}$*, the formula denotation function* D_I^g *is the total function from* D^Σ *to the power set of* U *such that*

for each $\tau \in T^\Sigma$*, for each* $\sigma \in S$*,*

$$\mathsf{D}_\mathsf{I}^g(\tau \sim \sigma) = \left\{ u \in \mathsf{U} \;\middle|\; \begin{array}{l} \mathsf{T}_\mathsf{I}^g(\tau)(u) \text{ is defined, and} \\ \mathsf{S}\left(\mathsf{T}_\mathsf{I}^g(\tau)(u)\right) \sqsubseteq \sigma \end{array} \right\},$$

for each $\tau_1, \tau_2 \in T^\Sigma$*,*

$$\mathsf{D}_\mathsf{I}^g(\tau_1 \approx \tau_2) = \left\{ u \in \mathsf{U} \;\middle|\; \begin{array}{l} \mathsf{T}_\mathsf{I}^g(\tau_1)(u) \text{ is defined,} \\ \mathsf{T}_\mathsf{I}^g(\tau_2)(u) \text{ is defined, and} \\ \mathsf{T}_\mathsf{I}^g(\tau_1)(u) = \mathsf{T}_\mathsf{I}^g(\tau_2)(u) \end{array} \right\},$$

for each $\rho \in R$*, for each* $x_1, \ldots, x_{Ar(\rho)} \in V$*,*

$$\mathsf{D}_\mathsf{I}^g\left(\rho(x_1, \ldots, x_{Ar(\rho)})\right) = \left\{ u \in \mathsf{U} \;\middle|\; \left\langle g(x_1), \ldots, g(x_{Ar(\rho)})\right\rangle \in \mathsf{R}(\rho) \right\},$$

for each $x \in V$*, for each* $\delta \in D^\Sigma$*,*

$$\mathsf{D}_\mathsf{I}^g\left(\exists x\delta\right) = \left\{ u \in \mathsf{U} \;\middle|\; \frac{\text{for some } u' \in \mathsf{U}}{u \in \mathsf{D}_\mathsf{I}^{g[x \mapsto u']}(\delta)} \right\},$$

for each $x \in V$*, for each* $\delta \in D^\Sigma$*,*

$$\mathsf{D}_\mathsf{I}^g\left(\forall x\delta\right) = \left\{ u \in \mathsf{U} \;\middle|\; \frac{\text{for each } u' \in \mathsf{U}}{u \in \mathsf{D}_\mathsf{I}^{g[x \mapsto u']}(\delta)} \right\},$$

for each $\delta \in D^\Sigma$*,* $\mathsf{D}_\mathsf{I}^g(\neg\delta) = \mathsf{U} \backslash \mathsf{D}_\mathsf{I}^g(\delta)$*,*

for each $\delta_1, \delta_2 \in D^{\Sigma}$, $\mathsf{D}_{\mathsf{I}}^{g}((\delta_1 \wedge \delta_2)) = \mathsf{D}_{\mathsf{I}}^{g}(\delta_1) \cap \mathsf{D}_{\mathsf{I}}^{g}(\delta_2)$
for each $\delta_1, \delta_2 \in D^{\Sigma}$, $\mathsf{D}_{\mathsf{I}}^{g}((\delta_1 \vee \delta_2)) = \mathsf{D}_{\mathsf{I}}^{g}(\delta_1) \cup \mathsf{D}_{\mathsf{I}}^{g}(\delta_2)$
for each $\delta_1, \delta_2 \in D^{\Sigma}$, $\mathsf{D}_{\mathsf{I}}^{g}((\delta_1 \rightarrow \delta_2)) = \big(\mathsf{U} \backslash \mathsf{D}_{\mathsf{I}}^{g}(\delta_1)\big) \cup \mathsf{D}_{\mathsf{I}}^{g}(\delta_2)$, *and*
for each $\delta_1, \delta_2 \in D^{\Sigma}$,
$\mathsf{D}_{\mathsf{I}}^{g}((\delta_1 \leftrightarrow \delta_2)) = ((\mathsf{U} \backslash \mathsf{D}_{\mathsf{I}}^{g}(\delta_1)) \cap (\mathsf{U} \backslash \mathsf{D}_{\mathsf{I}}^{g}(\delta_2))) \cup (\mathsf{D}_{\mathsf{I}}^{g}(\delta_1) \cap \mathsf{D}_{\mathsf{I}}^{g}(\delta_2))$.

Finally, the definition of *description denotation* is standard (Richter 2004: 177):

Definition 8 (description denotation) *For each signature* Σ*, for each* Σ *interpretation* $\mathsf{I} = \langle \mathsf{U}, \mathsf{r}, \mathsf{S}, \mathsf{A}, \mathsf{R} \rangle$*, the description denotation function* D_{I} *is the total function from* D_0^{Σ} *to the power set of* U *such that*
$\mathsf{D}_{\mathsf{I}}(\delta) = \big\{ u \in \mathsf{U} \,\big|\, \textit{for each } g \in \mathsf{G}_{\mathsf{I}},\ u \in \mathsf{D}_{\mathsf{I}}^{g}(\delta) \big\}$.

Also standard are the definitions of *grammar* and *model* (Richter 2004: 178–179), and it is not necessary (nor desirable) to define *exhaustive models* (Richter 2004: 179–180).

B Höhle's Problem – Comprehensive Solution

The comprehensive solution Höhle's Problem is similar to the simpler solution in Section 3 in the sense that it relies on the General Identity Principle in (12) and makes certain structures exempt from GIP by making them "escape" the definition of `are-copies`. But the structures that are exempt from GIP should be defined explicitly, rather than assuming that all cyclic structures are exempt. So the first step is to redefine `are-copies` in such a way that also cyclic structures which look the same are token identical, thus closing the loophole on which the simpler solution relies, and the second step is to explicitly define a new loophole allowing certain structures to "escape" the redefined `are-copies`.

In the first step, the relation `are-copies` is redefined so that it is sensitive to cycles. For this to work, it must have two more arguments which serve as the memory of objects through which the "currently examined" objects were reached. For example, the two new arguments are empty lists in the "top-level call" `are-copies`$(\langle\,\rangle, \boxed{1}, \langle\,\rangle, \boxed{2})$, which is used in the modified GIP in (25).

(25) $\forall\boxed{1}\forall\boxed{2}\,(\texttt{are-copies}(\langle\,\rangle, \boxed{1}, \langle\,\rangle, \boxed{2}) \rightarrow \boxed{1} \approx \boxed{2})$

However, when the sequences of objects visited on the paths to $\boxed{1}$ and $\boxed{2}$ are not empty, as for example in `are-copies`$(\langle \boxed{1_3}, \boxed{1_2}, \boxed{1_1} \rangle, \boxed{1}, \langle \boxed{2_3}, \boxed{2_2}, \boxed{2_1} \rangle, \boxed{2})$, it is possible that $\boxed{1}$ and $\boxed{2}$ are already in such sequences, i.e., it is possible that they have already been examined; e.g., in the example at hand, it might be the case that $\boxed{1} = \boxed{1_2}$ and $\boxed{2} = \boxed{2_2}$, i.e., that there are cycles $\boxed{1_2} \rightleftarrows \boxed{1_3}$ (i.e., $\boxed{1} \rightleftarrows \boxed{1_3}$) and $\boxed{2_2} \rightleftarrows \boxed{2_3}$ (i.e., $\boxed{2} \rightleftarrows \boxed{2_3}$). If so, it does not make sense to ask whether $\boxed{1}$ and $\boxed{2}$ are copies, as this question was already asked about them

(i.e., about $\boxed{1_2}$ and $\boxed{2_2}$) before; if all the other conditions on $\boxed{1}$ and $\boxed{2}$ being copies are satisfied, then it should be assumed that $\boxed{1}$ and $\boxed{2}$ indeed are copies. An ancillary relation, `member2`, is used to discover such cycles:

(26) $\texttt{member2}(\boxed{1}, \langle \boxed{\text{L}1h} | \boxed{\text{L}1t} \rangle, \boxed{2}, \langle \boxed{\text{L}2h} | \boxed{\text{L}2t} \rangle) \leftrightarrow$
$(\boxed{1} \approx \boxed{\text{L}1h} \wedge \boxed{2} \approx \boxed{\text{L}2h}) \vee \texttt{member2}(\boxed{1}, \boxed{\text{L}1t}, \boxed{2}, \boxed{\text{L}2t})$

In the running example, $\texttt{member2}(\boxed{1}, \langle \boxed{1_3}, \boxed{1_2}, \boxed{1_1} \rangle, \boxed{2}, \langle \boxed{2_3}, \boxed{2_2}, \boxed{2_1} \rangle)$ is true only when $\boxed{1}$ and $\boxed{2}$ are *parallel* members of the two corresponding lists, i.e., either $\boxed{1} = \boxed{1_1}$ and $\boxed{2} = \boxed{2_1}$, or $\boxed{1} = \boxed{1_2}$ and $\boxed{2} = \boxed{2_2}$, $\boxed{1} = \boxed{1_3}$ and $\boxed{2} = \boxed{2_3}$ (but not, e.g., when only $\boxed{1} = \boxed{1_2}$ and $\boxed{2} = \boxed{2_3}$). This makes sure that if cycles are discovered, they are of the same length in both structures.

With this ancillary relation in hand, the new `are-copies` relation, which closes the cyclicity loophole, is defined as in (27):

(27) $\forall \boxed{1} \forall \boxed{2}\ \texttt{are-copies}(\boxed{\text{L}1}, \boxed{1}, \boxed{\text{L}2}, \boxed{2}) \leftrightarrow$
$\texttt{member2}(\boxed{1}, \boxed{\text{L}1}, \boxed{2}, \boxed{\text{L}2}) \vee$
$\bigvee_{\sigma \in \mathcal{S}} (\boxed{1} \sim \sigma \wedge \boxed{2} \sim \sigma) \wedge$
$\bigwedge_{\alpha \in \mathcal{A}} (\boxed{1}\alpha \approx \boxed{1}\alpha \rightarrow \texttt{are-copies}(\langle \boxed{1} | \boxed{\text{L}1} \rangle, \boxed{1}\alpha, \langle \boxed{2} | \boxed{\text{L}2} \rangle, \boxed{2}\alpha))$

According to this definition, if $\boxed{1}$ and $\boxed{2}$ form parallel cycles, they are potential copies, and otherwise `are-copies` behaves as before: it checks the identity of species and whether values of all corresponding attributes are copies.

The combination of the new GIP in (25) and the new `are-copies` in (27) is exceptionless, so it is too strong – it is incompatible with the standard binding theory and the works mentioned in fn. 5. The following modified definition of `are-copies` makes it possible to specify exceptions: any objects satisfying the relation `int`, e.g., objects of sort *ref* (see (28)), will be exempted from GIP:

(28) $\forall \boxed{1}\ \texttt{int}(\boxed{1}) \leftrightarrow \boxed{1} \sim \mathit{ref}$

(29) $\forall \boxed{1} \forall \boxed{2}\ \texttt{are-copies}(\boxed{\text{L}1}, \boxed{1}, \boxed{\text{L}2}, \boxed{2}) \leftrightarrow$
$\texttt{member2}(\boxed{1}, \boxed{\text{L}1}, \boxed{2}, \boxed{\text{L}2}) \vee$
$\bigvee_{\sigma \in \mathcal{S}} (\boxed{1} \sim \sigma \wedge \boxed{2} \sim \sigma) \wedge$
$\bigwedge_{\alpha \in \mathcal{A}} (\boxed{1}\alpha \approx \boxed{1}\alpha \rightarrow \texttt{are-copies}(\langle \boxed{1} | \boxed{\text{L}1} \rangle, \boxed{1}\alpha, \langle \boxed{2} | \boxed{\text{L}2} \rangle, \boxed{2}\alpha)) \wedge$
$(\texttt{int}(\boxed{1}) \rightarrow \texttt{are-copies}(\boxed{\text{L}1}, \boxed{1}, \boxed{\text{L}2}, \boxed{2}))$

The way this works is, in short, as follows: if there are no parallel cycles (so $\texttt{member2}(\boxed{1}, \boxed{\text{L}1}, \boxed{2}, \boxed{\text{L}2})$ in the second line of (28) is false) and the usual conditions on what it means to be copies in the third and fourth line are satisfied, so $\boxed{1}$ and $\boxed{2}$ look the same, then either $\boxed{1}$ and $\boxed{2}$ are not specified as intensional ($\texttt{int}(\boxed{1})$ is false), in which case $\boxed{1}$ and $\boxed{2}$ are in the `are-copies`

relation, or they are specified as intensional (`int`(1) is true), in which case the above definition says that `are-copies`(L1, 1, L2, 2) (in the first line) is true iff `are-copies`(L1, 1, L2, 2) (in the last line) is true, so it is undetermined whether 1 and 2 are copies and, hence, they are not in the scope of GIP.

References

Abeillé, Anne & Rui Chaves. 2021. Coordination. In Müller et al. (2021). https://langsci-press.org/catalog/book/259. Forthcoming.

Bayer, Samuel. 1996. The coordination of unlike categories. *Language* 72(3). 579–616.

Beavers, John & Ivan A. Sag. 2004. Coordinate ellipsis and apparent non-constituent coordination. In Müller (2004) 48–69. http://csli-publications.stanford.edu/HPSG/2004/.

Blackburn, Patrick, Maarten de Rijke & Yde Venema. 2010. *Modal logic*. Cambridge: Cambridge University Press. Fourth printing (with corrections).

Chaves, Rui P. 2006. Coordination of unlikes without unlike categories. In Stefan Müller (ed.), *Proceedings of the HPSG 2006 conference*, 102–122. Stanford, CA: CSLI Publications.

Chaves, Rui P. 2008. Linearization-based word-part ellipsis. *Linguistics and Philosophy* 31. 261–307.

Chaves, Rui P. 2013. Grammatical alignments and the gradience of lexical categories. In Philip Hofmeister & Elisabeth Norcliffe (eds.), *The core and the periphery: Data-driven perspectives on syntax inspired by Ivan A. Sag*, 167–220. Stanford, CA: CSLI Publications.

Crysmann, Berthold. 2003. An asymmetric theory of peripheral sharing in HPSG: Conjunction reduction and coordination of unlikes. In Gerhard Jäger, Paola Monachesi, Gerald Penn & Shuly Wintner (eds.), *Proceedings of Formal Grammar 2003*, 47–62. http://cs.haifa.ac.il/~shuly/fg03/.

Dalrymple, Mary. 2017. Unlike phrase structure category coordination. In Victoria Rosén & Koenraad De Smedt (eds.), *The very model of a modern linguist*, vol. 8 Bergen Language and Linguistics Studies, 33–55. Bergen: University of Bergen Library. doi:http://dx.doi.org/10.15845/bells.v8i1. https://bells.uib.no/index.php/bells/issue/view/249.

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

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning and use of English interrogatives*. Stanford, CA: CSLI Publications.

Höhle, Tilman N. 1999. An architecture for phonology. In Robert D. Borsley & Adam Przepiórkowski (eds.), *Slavic in Head-driven Phrase Structure*

Grammar, 61–90. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/site/1575861747.shtml. Reprinted in Müller et al. 2019.

Kepser, Stephan. 2004. On the complexity of RSRL. In Lawrence S. Moss & Richard T. Oehrle (eds.), *Proceedings of the joint meeting of the 6th Conference on Formal Grammar and the 7th Conference on Mathematics of Language (10–12 august 2001, helsinki, finland)*, 146–162.

King, Paul John. 1999. Towards truth in HPSG. In Valia Kordoni (ed.), *Tübingen studies in Head-driven Phrase Structure Grammar* Arbeitspapiere des Sonderforschungsbereichs 340, Bericht Nr. 132, 301–352. Tübingen: Universität Tübingen.

Levine, Robert D. 2011. Linearization and its discontents. In Stefan Müller (ed.), *Proceedings of the HPSG 2011 conference*, 126–146. Stanford, CA: CSLI Publications.

Levy, Roger. 2001. Feature indeterminacy and the coordination of unlikes in a totally well-typed HPSG. Unpublished manuscript, http://idiom.ucsd.edu/~rlevy/papers/feature-indet.pdf.

Meurers, Walt Detmar. 1998. Raising spirits (and assigning them case). Handout for a talk delivered during the Groningen-Tübingen Kolloquium, Tübingen, 27th November 1998.

Müller, Stefan (ed.). 2004. *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar, Center for Computational Linguistics, Katholieke Universiteit Leuven*. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2004/.

Müller, Stefan, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.). 2021. *Head-driven Phrase Structure Grammar: The handbook*. Berlin: Language Science Press. https://langsci-press.org/catalog/book/259. Forthcoming.

Müller, Stefan, Marga Reis & Frank Richter (eds.). 2019. *Beiträge zur deutschen Grammatik: Gesammelte Schriften von Tilman N. Höhle*. Berlin: Language Science Press.

Patejuk, Agnieszka & Adam Przepiórkowski. 2021. Category mismatches in coordination vindicated. *Linguistic Inquiry* doi:10.1162/ling_a_00438. Forthcoming.

Pollard, Carl. 1999. Strong generative capacity in HPSG. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and constructional aspects of linguistic explanation*, 281–297. Stanford, CA: CSLI Publications.

Pollard, Carl. 2001. Cleaning the HPSG garage: Some problems and some proposals. Unpublished manuscript, Ohio State University; http://utkl.ff.cuni.cz/~rosen/public/cp_garage.ps.

Pollard, Carl. 2014. Type-logical HPSG. In Gerhard Jäger, Paola Monachesi, Gerald Penn & Shuly Wintner (eds.), *Proceedings of Formal Grammar 2004*, 111–128. CSLI Publications. https://web.stanford.edu/group/

cslipublications/cslipublications/FG/2004/fg-2004-paper08.pdf.

Pollard, Carl & Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar*. Chicago, IL: Chicago University Press / CSLI Publications.

Przepiórkowski, Adam. 1999. *Case assignment and the complement-adjunct dichotomy: A non-configurational constraint-based approach*: Universität Tübingen Ph.D. dissertation.

Przepiórkowski, Adam. 2021. LFG and Head-driven Phrase Structure Grammar. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*, Berlin: Language Science Press. https://langsci-press.org/catalog/book/312. Forthcoming.

Przepiórkowski, Adam & Agnieszka Patejuk. 2012. On case assignment and the coordination of unlikes: The limits of distributive features. In Miriam Butt & Tracy Holloway King (eds.), *The proceedings of the LFG'12 conference*, 479–489. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/LFG/17/lfg12.html.

Przepiórkowski, Adam & Agnieszka Patejuk. 2021. Coordinate structures without syntactic categories. In I Wayan Arka, Ash Asudeh & Tracy Holloway King (eds.), *Modular design of grammar: Linguistics on the edge*, 205–220. Oxford: Oxford University Press. https://global.oup.com/academic/product/modular-design-of-grammar-9780192844842. Forthcoming.

Richter, Frank. 2004. *A mathematical formalism for linguistic theories with an application in Head-driven Phrase Structure Grammar*: Universität Tübingen Ph.D. dissertation (defended in 2000).

Richter, Frank. 2007. Closer to the truth: A new model theory for HPSG. In James Rogers & Stephan Kepser (eds.), *Model-theoretic syntax at 10*, 99–108.

Richter, Frank. 2021. Formal background. In Müller et al. (2021). https://langsci-press.org/catalog/book/259. Forthcoming.

Sailer, Manfred. 2003. *Combinatorial semantics and idiomatic expressions in Head-driven Phrase Structure Grammar*: Universität Tübingen Ph.D. dissertation (defended in 2000).

Søgaard, Anders & Martin Lange. 2009. Polyadic dynamic logics for HPSG parsing. *Journal of Logic, Language and Information* 18. 159–198.

Williams, Edwin. 1981. Transformationless grammar. *Linguistic Inquiry* 12(4). 645–653.

Yatabe, Shûichi. 2004. A comprehensive theory of coordination of unlikes. In Müller (2004) 335–355. http://csli-publications.stanford.edu/HPSG/2004/.

Yatabe, Shûichi & Wai Lok Tam. 2021. In defense of an HPSG-based theory of non-constituent coordination: A reply to Kubota and Levine. *Linguistics and Philosophy* 44. 1–77.