Head-Driven Phrase Structure Grammar as a theory of the syntax-semantics interface

Shûichi Yatabe
University of Tokyo

September 2, 2024

1 Introduction

This document is partly a tutorial and partly a polemic about the theory of the syntax-semantics interface advocated in Yatabe and Tam (2021), which achieves descriptive adequacy in empirical domains in which other theories have persistently fallen short and does so in a way that calls into question some of the widely held fundamental assumptions about the syntax-semantics interface.

The theory that I will describe in this document has achieved, using only simple and well-motivated mechanisms, a reasonably comprehensive descriptive adequacy in regard to phenomena such as right-node raising, left-node raising, and split-antecedent relative clauses, none of which has received satisfactory treatment in other theories of grammar, as I will demonstrate below. This achievement is notable because while it is true that descriptive adequacy is not the most important metric by which to evaluate linguistic theories, a theory that is not even descriptively adequate cannot rationally be expected to attain any sort of explanatory adequacy.

As I will explain in detail below, the theory presented in Yatabe and Tam (2021) makes empirical gains not by grafting an additional mechanism onto some previous theory but by modifying some of the fundamental assumptions about the syntax-semantics interface shared by most previous theories of grammar. Specifically, the theory is based on the view that semantic composition is performed each time a larger *prosodic* constituent is created, not each time a larger syntactic constituent is created.

In what follows, I will first delineate the standard theory of the syntax-semantics interface, according to which semantic composition is mostly a series of function application that is performed each time a larger syntactic constituent is created, before presenting what I believe to be the better theory and comparing it with the standard theory. No previous familiarity with theories of the syntax-semantics interface is presupposed, although some familiarity with theories of syntax is.

2 The standard theory of semantic composition

While the meanings of lexical items tend to be quite nebulous and elusive, the way the meanings of lexical items are combined with each other to yield the meaning of a sentence is arguably rule-governed. The theory depicted in Heim and Kratzer (1998) has been the standard theory of that rule-governed aspect of linguistic meaning.

In this theory, which incorporates the so-called T-model, semantic interpretation of a sentence is performed on the basis of its LF representation. In the case of a simple sentence like *Chris talks*, its LF representation is identical to its surface-level syntactic structure. A proper noun like *Chris* is assumed to denote a specific entity like an actual person named

Chris, whom we might decide to refer to using the symbol "c". In accordance with the general assumption that an intransitive verb denotes a function that maps an entity either to 1 or to 0, the intransitive verb *talks* is assumed to denote a function that maps an entity to 1 if that entity talks and to 0 if it does not talk. Let us refer to that specific function using the symbol "talks′". The denotation of a phrase consisting of two daughter nodes is computed in most cases by applying the function denoted by one of the daughters to the denotation of the other daughter. In the example above, the denotation of the sentence is computed by applying the function denoted by *talks* to the denotation of *Chris*. The result of that function application is talks′(c), which equals 1 if Chris talks and 0 if he or she does not. Note that a formula of the form "$f(x)$" expresses the value that the function f returns when it is applied to the argument x; the formula above expresses the value that the function *talks′* returns when it is applied to the argument c. Relying on the convention of using the notation $[\![E]\!]$ to express the denotation of a linguistic expression E, what we have done so far can be summarized as follows.

(1) $[\![\text{Chris talks}]\!]$
$= [\![\text{talks}]\!]\,(\,[\![\text{Chris}]\!]\,)$
$= \text{talks}'(c)$
$= \begin{cases} 1 & \text{if Chris talks} \\ 0 & \text{otherwise} \end{cases}$

Given the convention that 1 means true and 0 means false, we can say that the procedure depicted here has computed the truth conditions of the sentence *Chris talks*, that is, the conditions that the world has to satisfy in order for the sentence to be true. While the theory says virtually nothing about the meanings of lexical items, it does state how they are combined with each other to yield the sentence meaning.

Let us next consider the meaning of the sentence *Chris speaks Russian*, which consists of a transitive verb and its object and subject. We will use the symbol r to refer to the entity that the word *Russian* denotes. If the denotation of this sentence is to be computed, as in the previous example, by applying the function denoted by the verb phrase to the denotation of the subject, the verb phrase *speaks Russian* in this sentence has to denote the function that maps an entity to 1 if it speaks Russian and to 0 if it does not. Let us call this function f_r. More generally, we might use the notation f_x to refer to the function that maps an entity to 1 if it speaks x and to 0 if it does not. Now, if we are to perform as much semantic composition as possible through function application, we need to assume that the denotation of the verb phrase is computed by applying the function denoted the transitive verb to the denotation of its object. This means that the denotation of the verb *speaks* has to be that function that maps an entity x to f_x. We will refer to this function using the symbol "speaks′". Summarizing, the denotation of the sentence in question is computed as in (2) in this theory.

(2) $[\![\,[_{\text{S}}\ \text{Chris}\ [_{\text{VP}}\ \text{speaks Russian}]]\,]\!]$
$= [\![\,[_{\text{VP}}\ \text{speaks Russian}]\,]\!]\,(\,[\![\text{Chris}]\!]\,)$
$= (\,[\![\text{speaks}]\!]\,(\,[\![\text{Russian}]\!]\,))(\,[\![\text{Chris}]\!]\,)$
$= (\text{speaks}'(r))(c)$
$= f_r(c)$
$= \begin{cases} 1 & \text{if Chris speaks Russian} \\ 0 & \text{otherwise} \end{cases}$

The things that are postulated as the denotations of expressions are categorized into types. Entities like a person named Chris and the Russian language are said to be of type e. The things that sentences denote, namely 0 and 1, are called truth values and are said to be of type t. Functions that map things of type x to things of type y are said to be of type $\langle x, y\rangle$. Thus, the denotation of an intransitive verb is of type $\langle e, t\rangle$ and the denotation of a transitive verb is of type $\langle e, \langle e, t\rangle\rangle$.

We next consider the meaning of the sentence *Chris speaks every language*, whose LF representation is shown in (3).

(3) $[[\text{every language}]_i\ [_{\text{S}}\ \text{Chris}\ [_{\text{VP}}\ \text{speaks}\ t_i]]]$

The denotation of the noun *language*, which we will write as "language′", is assumed to be a function of type $\langle e, t\rangle$ that maps an entity x to 1 if and only if x is a language. The denotation of the S out of which the expression *every language* has been extracted is assumed to be the function of type $\langle e, t\rangle$ that maps an entity x to the denotation of the sentence "Chris speaks X", where X is a DP whose denotation is x. Following the convention of lambda calculus, let us use a formula of the form $\lambda x[R]$ to express that function that returns the value R when x is given as its argument. Then the denotation of the S in question can be written as in (4).

(4) $\lambda x[(\text{speaks}'(x))(c)]$

The denotation of the determiner *every*, which we will write as "every′", is assumed to be the function of type $\langle\langle e, t\rangle, \langle\langle e, t\rangle, t\rangle\rangle$ defined in (5).

(5) For any two functions P and Q of type $\langle e, t\rangle$, $(\text{every}'(P))(Q) = 1$ if and only if for every entity y such that $P(y) = 1$, $Q(y) = 1$.

The denotation of the DP *every language* and the denotation of the entire sentence are both assumed to be computed via function application. Given these assumptions, the LF representation in (3) is interpreted as in (6).

(6)
$$\begin{aligned}&[\![[[\text{every language}]_i\ [_{\text{S}}\ \text{Chris}\ [_{\text{VP}}\ \text{speaks}\ t_i]]]]\!]\\ &= (\text{every}'(\text{language}'))(\lambda x[(\text{speaks}'(x))(c)])\\ &= \begin{cases} 1 & \text{if for every entity } y \text{ such that language}'(y) = 1,\\ & \quad \lambda x[(\text{speaks}'(x))(c)](y) = 1\\ 0 & \text{otherwise}\end{cases}\end{aligned}$$

Let us see what the truth conditions computed here amount to. First, $\text{language}'(y) = 1$ if and only if y is a language. Second, $\lambda x[(\text{speaks}'(x))(c)](y) = (\text{speaks}'(y))(c) = 1$ if and only if Chris speaks y. Thus, the calculation in (6) is saying, correctly, that the sentence *Chris speaks every language*, whose LF representation is (3), is true if and only if for every entity y such that y is a language, Chris speaks y.

This, in a nutshell, is how the denotations of sentences are computed on the basis of their LF representations in the standard theory.

3 Some unnatural features of the standard theory

The standard theory of the syntax-semantics interface has some unnatural features as a consequence of its reliance on the mechanism of function application. Consider first the sentences in (7), where the word *or* appears to express something akin to logical disjunction.

(7) a. Chris walks or Pat talks.
b. Chris walks or talks.
c. Chris reads or speaks every language.

Intuitively, the word *or* has the same meaning in these three sentences, but supposing that a coordinate structure has a binary-branching structure of the form [Disjunct1 [or Disjunct2]] and supposing that semantic composition is performed by function application here as well, the denotation of *or* has to be of type $\langle t, \langle t, t\rangle\rangle$ in (7a), $\langle\langle e, t\rangle, \langle\langle e, t\rangle, \langle e, t\rangle\rangle\rangle$ in (7b), and $\langle\langle e, \langle e, t\rangle\rangle, \langle\langle e, \langle e, t\rangle\rangle, \langle e, \langle e, t\rangle\rangle\rangle\rangle$ in (7c). Specifically, the denotation of *or* has to be (8a) in (7a), (8b) in (7b), and (8c) in (7c). Here, I am using the symbol "or′" to express the function that maps a pair of truth values $\langle x, y\rangle$ to 0 if $x = y = 0$ and to 1 otherwise.

(8) a. $\lambda v[\lambda u[\,\text{or}'(u, v)]]$
b. $\lambda g[\lambda f[\lambda x[\,\text{or}'(f(x), g(x))]]]$
c. $\lambda q[\lambda p[\lambda y[\lambda x[\,\text{or}'((p(y))(x), (q(y))(x))]]]]$

It has been noted in the literature (Gazdar (1980) and Partee and Rooth (1983)) that the denotations like those in (8) can be unified and that it is therefore possible to have a single lexical entry for a word like *or* that can be used irrespective of the syntactic category of the constituents being coordinated. In the version of this type of analysis proposed in Partee and Rooth (1983), the denotations of words like *and* and *or* that appear to mean something like logical conjunction and disjunction are expressed by the operators $\sqcap$ and $\sqcup$, which are defined as in (9). Note that functions are represented here as sets of ordered pairs. Note also that I use the symbol "and′" to express the function that maps a pair of truth values $\langle x, y\rangle$ to 1 if $x = y = 1$ and to 0 otherwise.

(9) Pointwise definition of $\sqcap$ and $\sqcup$

a. $X \sqcap Y$

$$= \begin{cases} \text{and}'(X, Y) \text{ if } X \text{ and } Y \text{ are truth values} \\ \{\langle z, x \sqcap y\rangle : \langle z, x\rangle \in X \text{ and } \langle z, y\rangle \in Y\} \text{ if } X \text{ and } Y \text{ are functions} \end{cases}$$

b. $X \sqcup Y$

$$= \begin{cases} \text{or}'(X, Y) \text{ if } X \text{ and } Y \text{ are truth values} \\ \{\langle z, x \sqcup y\rangle : \langle z, x\rangle \in X \text{ and } \langle z, y\rangle \in Y\} \text{ if } X \text{ and } Y \text{ are functions} \end{cases}$$

The fact that such unified lexical entries are *possible*, however, resolves only part of the problem. As things stand, it is not *required* that words like *and* and *or* be given the denotations defined in (9). Nothing in the theory prevents there from being a lexical entry whose denotation is, say, (8b), a denotation that allows the lexical entry to function as a coordinator only when the denotations of the constituents being coordinated are of type $\langle e, t\rangle$. If there is not any such morpheme in any language, then that will remain an unexplained fact.

The situation is analogous with regard to nominal conjunction that is used to express groups. Consider the following examples.

(10) the Alps and the Himalayas

(11) ten men and women (from Bayer (1997))

Example (10) can refer to a group of mountains such that some members of that group are the Alps and the rest are the Himalayas; the conjunction word *and* seems to be forming an expression that denotes a group consisting of the denotations of the two conjuncts. The word *and* in example (11) appears to have the same function. Example (11) can be used to refer to a group of ten people such that some members of the group are men and the rest are women, when the word *and* is taken to conjoin two NPs, *men* and *women*, rather than two DPs, *ten men* and *women*. Intuitively, the word *and* here also appears to be constructing an expression that denotes a group consisting of the denotations of the two conjuncts. According to the standard theory, however, while the word *and* in (10) is conjoining two expressions whose denotations are of type e and thus can be viewed as having a denotation that creates a group out of two entities, the coordinator in (11) is conjoining two expressions whose denotations are not entities but functions of type $\langle e, t\rangle$ and thus has to be given a different denotation. It is therefore predicted, implausibly, that there can be languages that use different morphemes for DP-level group-forming conjunction and NP-level group-forming conjunction. It has been noted in the literature (see Bayer (1997) and the references cited there) that the two denotations needed for DP-level and NP-level group-forming conjunction can be unified with each other, but just as in the case of logical conjunction and disjunction, the fact that such unification is *possible* does not make the potential problem go away altogether.

What we have seen in this section in no way invalidates the standard theory, but it provides a reason to suspect that there may be a better theory.

4 Minimal Recursion Semantics

Minimal Recursion Semantics (MRS), proposed in Copestake et al. (2005), is one of the semantic formalisms that do not share the unnaturalness of the standard theory noted in section 3. In this theory, the meaning of a predicate and the meaning of its arguments are combined with each other not through function application but by using a mechanism that was first used in the Discourse Representation Theory (DRT), proposed in Kamp (1984). I will outline in this section a slightly modified version of MRS as a first step in explaining the theory presented in Yatabe and Tam (2021).

In MRS, the meaning of a word is typically represented by what is called an elementary predication, which consists of a predicate and all its arguments, and the meaning of a phrase is represented by a sequence of elementary predications. Each elementary predication in such a sequence is given a handle, which indicates precisely how that elementary predication is to fit into the overall semantic representation. Let us take the following sequence of three elementary predications as our first example.

(12) $\langle h0 : \mathrm{identical}(x, \mathrm{Chris}),\ h0 : \mathrm{speaks}(x, y),\ h0 : \mathrm{identical}(y, \mathrm{Russian})\rangle$

These three elementary predications, each prefaced with its handle, namely $h0$, represent the meaning of the words *Chris*, *speaks*, and *Russian* respectively, and the sequence as a whole is meant to represent the meaning of the sentence *Chris speaks Russian*. An elementary predication of the form "identical(i, j)" is assumed to be true if and only the denotation of the first argument is identical to that of the second argument. An elementary predication of the form "speaks(x, y)" is assumed to be true if and only if the denotation of x speaks the denotation of y. In an MRS representation, elementary predications that are prefaced by the same handle are interpreted as being conjoined with each other. The variables that are not bound by any quantifier are interpreted as bound by an unexpressed existential quantifier at the global top level. Therefore the MRS representation in (12) as a whole is true if and only if there are entities x and y such that x is Chris, y is the Russian language, and x speaks y.

Consider next the MRS representation in (13), which illustrates how quantifiers are handled.

(13) $\langle\ h0 : \mathrm{every}(x, h1, h2),\ h1 : \mathrm{smart}(x),\ h1 : \mathrm{student}(x),\ h3 : \mathrm{probably}(h4),$
$h5 : \mathrm{agrees}(x)\ \rangle$

The five elementary predications in (13) represent the meaning of the words *every*, *smart*, *student*, *probably*, and *agrees* respectively, and the sequence as a whole is meant to represent the two possible readings of the sentence *Every smart student probably agrees*, one in which *every smart student* outscopes *probably* and another in which the universal quantifier is outscoped by the adverb. The three arguments of the predicate *every* are the variable that the quantifier binds, the restrictor of the quantifier, and the (nuclear) scope of the quantifier, respectively. An elementary predication of the form "every(x, S, T)" is assumed to be true if and only if every possible value of x that makes S true makes T true as well. The predicates *smart*, *student*, and *agrees* are interpreted in the expected way. An elementary predication of the form "probably(S)" is assumed to be true if and only if the formula S is probably true.

When the handle of a set of elementary predications is identical to the handle filling an argument slot of a predicate, the set of elementary predications is interpreted as filling that argument slot. Thus, the second and the third elementary predication in this example, which both have the handle $h1$, are interpreted as filling the second argument slot of *every*, the argument slot that represents the restrictor of the quantifier.

An MRS representation can serve as an underspecified representation that stands for more than one semantic interpretation because it is assumed that even different handles (such as $h1$ and $h2$) can optionally be interpreted as pointing to the same location in the semantic representation. Let us use the notation $h \approx h'$ to indicate that two handles h and h' point to the same location in the semantic representation, and let us say that h and h' are *locationally equivalent* when $h \approx h'$ holds. I will use example (13) to illustrate what happens when some of the handles in an MRS representation are locationally equivalent. Suppose that $h2 \approx h3$ and $h4 \approx h5$ in (13). Then the elementary predication "probably($h4$)" becomes the nuclear scope of *every* and the elementary predication "agrees(x)" becomes the sole argument of *probably*, so the representation in (13) as a whole becomes equivalent to (14).

(14) $\text{every}(x, \text{smart}(x) \wedge \text{student}(x), \text{probably}(\text{agrees}(x)))$

On the other hand, the representation in (13) will become equivalent to (15) if $h4 \approx h0$ and $h2 \approx h5$.

(15) $\text{probably}(\text{every}(x, \text{smart}(x) \wedge \text{student}(x), \text{agrees}(x)))$

The former is the reading in which the universal quantifier takes wide scope over the adverbial, and the latter is the reading in which the adverbial takes scope over the universal quantifier.

A handle h is said to *immediately outscope* a handle h' if and only if h or a handle locationally equivalent to h is attached to an elementary predication that has an argument slot filled by h' or a handle locationally equivalent to h'. We say that *h outscopes h'* if and only if h either immediately outscopes h' or immediately outscopes a handle that outscopes h'. We will write $h \geq h'$ to mean that h either is locationally equivalent to or outscopes h'.

The two readings shown in (14) and (15) are the only readings expressed by the underspecified representation in (13) because it is assumed (i) that there is a global top handle that outscopes every other handle that is not locationally equivalent to it, (ii) that, for each handle filling an argument slot of a predicate, there must be at least one elementary predication whose handle is locationally equivalent to it, and (iii) that an elementary predication cannot be interpreted as filling more than one argument slot. The third assumption, which requires that elementary predications in an MRS representation should form a tree, prevents $h5$ from being locationally equivalent to $h1$ and to $h2$ at the same time in (13), for example. Another constraint that is imposed on MRS representations is that a variable that is bound by a quantifier in an MRS representation cannot be reused as a free variable or as a variable bound by another quantifier in the same MRS representation, although this condition, which is referred to as the variable-binding condition in the literature, plays no role in the above example.

MRS is easy to integrate into a grammatical theory like Head-Driven Phrase Structure Grammar (HPSG) (see Pollard and Sag (1994) and Sag, Wasow, and Bender (2003)).[1] In HPSG, each node in a syntactic phrase-structure tree is associated with a semantic representation that expresses the meaning of that node. In order to incorporate MRS into HPSG, we need to assume that the semantic representation associated with each node consists of the following:

(i) a list of elementary predications, L, corresponding to the semantic contribution that node makes to the meaning of the sentence as a whole,

(ii) the index, x,

---

[1] MRS is not the only theory of the syntax-semantics interface that has been pursued in combination with HPSG. See Koenig and Richter (2021) for an overview.

(iii) the local top handle, h, which is the lowest handle such that for each elementary predication in L, either the handle of that elementary predication is locationally equivalent to or outscoped by h or that elementary predication is part of a quantifier that binds a variable other than x and takes scope at a node higher than the node concerned,[2] and

(iv) the semantic head handle, h', which is the lowest handle such that for each elementary predication in L, either the handle of that elementary predication is locationally equivalent to or outscoped by h' or that elementary predication is part of a quantifier that binds a variable other than x.

If the root node of a sentence is associated with a semantic representation equivalent to (14) above, the local top handle and the semantic head handle of that node will be locationally equivalent to the handle attached to the *every* predicate and the handle attached to the *probably* predicate respectively. If the node is associated with a semantic representation equivalent to (15), the local top handle and the semantic head handle of the node will both be locationally equivalent to the handle attached to the *probably* predicate.

The lexical entries of expressions often indicate which handles and which indices are to be identical. The lexical entry for a proper name like *Chris* requires its own index to be identical to the first argument in the sole elementary predication it is associated with, namely the one whose predicate is *identical*. The lexical entry for *every* requires that the local top handle of its NP complement be identical to the second argument (i.e. the restrictor) of the elementary predication whose predicate is *every*. It also requires (i) that its own index be identical to the index of its complement NP and (ii) that its own index be identical to the first argument (i.e. the variable) of the elementary predication whose predicate is *every*. The lexical entry for *smart* requires that the index of the nominal modified by it be identical to the argument of the *smart* predicate. The lexical entry for a common noun like *student* requires its own index to be identical to the sole argument in the sole elementary predication that it is associated with. The lexical entry for *probably* requires the local top handle of the expression that it modifies (i.e. its sister) to be identical to the sole argument in the sole elementary predication that it is associated with. The lexical entry for *agrees* requires that the index of its subject be identical to the sole argument of the *agrees* predicate. And each lexical entry requires its own semantic head handle to be identical to the handle attached to one of the elementary predications associated with it.

There are grammatical principles that constrain the relationship between the semantic representation associated with a phrasal node and the semantic representations associated with its daughter nodes. The index of a headed phrase is required to be identical to the index of its head daughter; this is part of the effect of what is called the Semantic Inheritance Principle in Sag, Wasow, and Bender (2003). And the list of elementary predications associated with a phrase is the concatenation of the lists of elementary predications associated with its daughters; this is what is called the Semantic Compositionality Principle in Sag, Wasow, and Bender (2003).

The local top handle and the semantic head handle of a phrasal node are subject to the following constraints.[3]

(16) a. If a phrase consists of a nominal head daughter (i.e. a head daughter whose HEAD value is either of type *noun* or of type *det*) and its complement, the semantic head handle of the phrase must be identical to the local top handle of the head daughter.

---

[2] What is called the local top handle in Copestake et al. (2005) is the same as what is called the semantic head handle in the present document. What is called the local top handle in the present document is not given a name in Copestake et al. (2005).

[3] Some of the constraints in (16) are different from the corresponding constraints presented in Copestake et al. (2005).

b. If a phrase consists of a non-nominal head daughter and its argument (i.e. either a complement or a subject), the local top handle and the semantic head handle of the phrase must be identical to the local top handle and the semantic head handle of the head daughter respectively.

c. If a phrase consists of a head daughter and an adjunct whose denotation is not intersective, the semantic head handle of the phrase must be identical to the local top handle of the non-head daughter.

d. If a phrase consists of a head daughter and an adjunct whose denotation is intersective, the semantic head handle of the phrase, the local top handle of the non-head daughter, and the local top handle of the head daughter must be identical to each other.

The denotation of an adjunct is said to be *intersective* when the meaning of the phrase consisting of the adjunct and its sister can be viewed as a conjunction of the denotation of the adjunct and that of its sister. For instance, the adjective *smart* in the phrase *smart student* is intersective because being a smart student amounts to being smart and being a student at the same time. The adverb *probably*, on the other hand, is not intersective because it is not possible to paraphrase an expression consisting of *probably* and its sister by conjoining the meaning of that sister node with something else; the sentence *Probably it rained* is not synonymous with a sentence of the form *It rained and X*, irrespective of what *X* says, since the former does not entail that it rained whereas the latter does.

Additionally, when h and h' are the local top handle and the semantic head handle of a node, it is required (i) that h' should either be locationally equivalent to or be outscoped by h and (ii) that there is no handle h'' such that h'' is outscoped by h and outscopes h' and h'' is not locationally equivalent to the third argument of a quantificational elementary predication. I will refer to this requirement as the LTOP-over-SEMHEAD requirement below, using the names of the features that will be used to represent local top handles and semantic head handles. When the local top handle h and the semantic head handle h' of a node satisfy the LTOP-over-SEMHEAD requirement, I will represent that situation with the notation $h =_q h'$, following the convention used in Copestake et al. (2005) (although the way the relation represented by this notation is used here differs significantly from the way it is used in Copestake et al. (2005)).

Let us see how all the lexical entries and grammatical principles work together to ensure that the sentence *Every smart student probably agrees* is associated with an appropriate MRS representation. Figure 1 is the structure assigned to this sentence. Each node in the tree shown here is associated with a feature structure that has two features, PHON and SYNSEM. The value of the PHON feature shows how the node is to be pronounced, and the value of the SYNSEM feature shows the syntactic and semantic properties of the node. The value of the SYNSEM feature is a feature structure with two features, CAT and CONT, whose values show the syntactic properties and the semantic properties of the node respectively. The value of the CONT feature is a feature structure with features such as LTOP, SEMHEAD, INDEX, and EP. The LTOP value, the SEMHEAD value, and the EP value associated with a node is the local top handle of that node, the semantic head handle of that node, and the list of elementary predications associated with that node, respectively. The INDEX values associated with non-nominal nodes are suppressed in Figure 1 because they do not play any role here.

In HPSG, a phrase-structure tree is deemed grammatical if and only if each local subtree in it corresponds to one of the grammar rules (such as the Head-Adjunct Rule) and obeys all the grammatical principles. A local subtree in a given tree is a structure that consists of a non-leaf node in the tree and all its daughters. I will examine each local subtree contained in the tree shown in Figure 1.

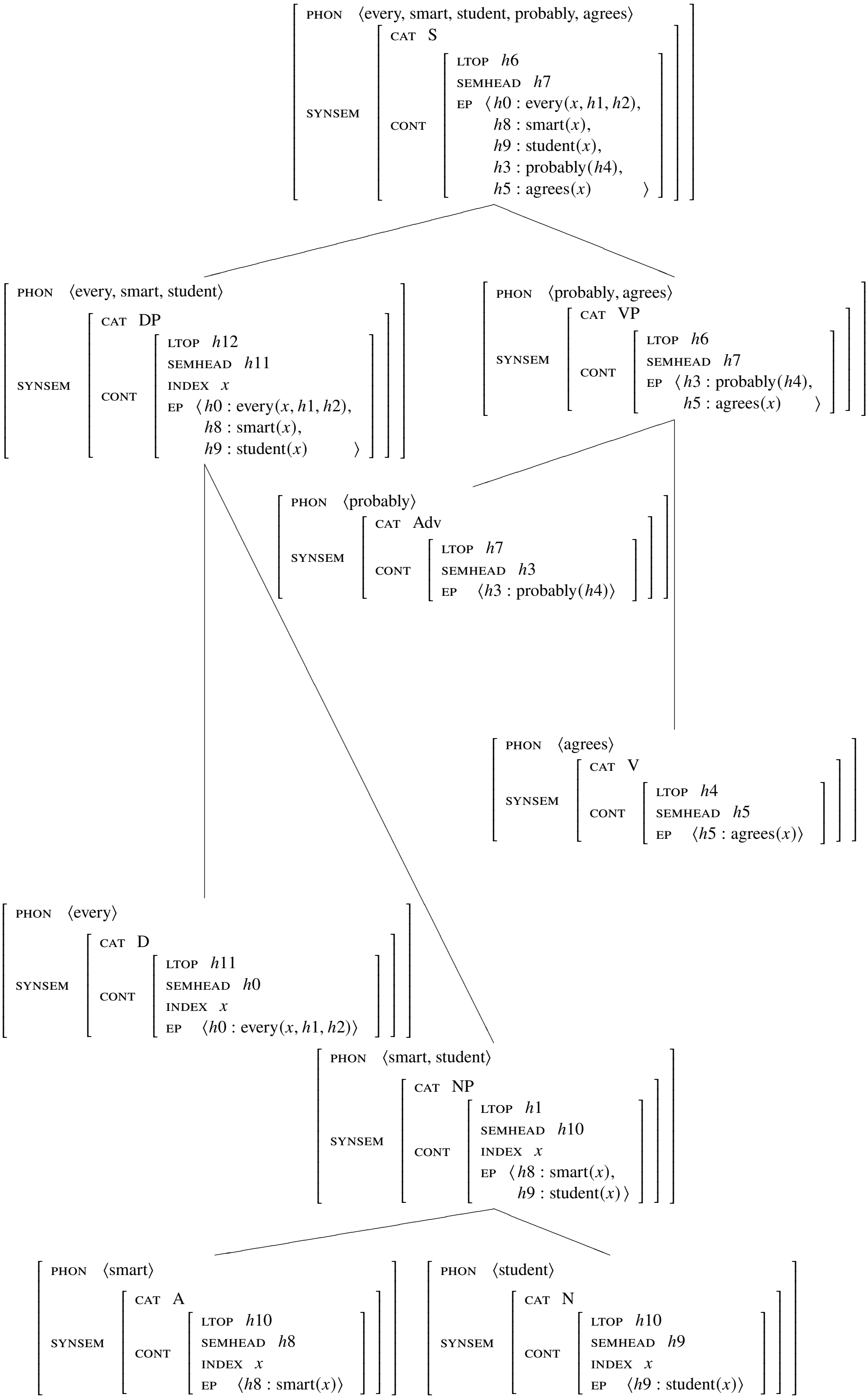

Figure 1: The structure of the sentence *Every smart student probably agrees*

In the local subtree consisting of the NP node and its two daughters, viz. the A node and the N node, the SEMHEAD value of the mother node and the LTOP values of the two daughter nodes are identified with each other due to (16d). The argument of the *smart* predicate is identified with the index of the N due to the lexical entry of *smart*, the argument of the *student* predicate is also identified with that index due to what is stated in the lexical entry of *student*, and the index of the NP is identified with that index as well due to the Semantic Inheritance Principle.

In the local subtree consisting of the DP node and its two daughters, viz. the D node and the NP node, the SEMHEAD value of the mother is identified with the LTOP value of the D node due to (16a). The second argument of the *every* predicate is identified with the LTOP value of the NP node due to what is stated in the lexical entry of *every*. The first argument of the *every* predicate is identified with the index of the NP as well as with the index of the D in accordance with the lexical entry of *every*, and the same index is shared by the DP node as well due to the Semantic Inheritance Principle.

In the local subtree consisting of the VP node and its two daughters, viz. the Adv node and the V node, the SEMHEAD value of the mother is identified with the LTOP value of the Adv node due to (16c). The sole argument of the *probably* predicate is identified with the LTOP value of the V node due to what is stated in the lexical entry of *probably*.

And in the topmost local subtree, in which the DP node and the VP node are immediately dominated by the S node, the LTOP value and the SEMHEAD value of the S node are identified with the LTOP value and the SEMHEAD value of the VP node due to (16b). The index of the subject is identified with the sole argument of the *agrees* predicate due to what is stated in the lexical entry of *agrees*. The EP value of the S node contains all the elementary predications associated with the leaf nodes due to the Semantic Compositionality Principle.

Because of the LTOP-over-SEMHEAD requirement, we know the following: $h6 =_q h7, h12 =_q h11, h11 =_q h0, h1 =_q h10, h10 =_q h8, h10 =_q h9, h7 =_q h3$, and $h4 =_q h5$. At the same time, we know that $h3 \geq h4$ because one of the elementary predications is of the form "$h3$: probably($h4$)". Therefore $h3 \geq h5$; in other words, the elementary predication whose predicate is *agrees* must be in the scope of *probably*. This means that $h3$ is not locationally equivalent to or outscoped by $h1$; if it were, then $h5$ would also be either locationally equivalent to or outscoped by $h1$ and there would be no elementary predication left that could fill the third argument slot of the *every* predicate, i.e. the slot represented by $h2$, because from the above we know that $h1 \geq h10 \geq h8$ and that $h1 \geq h10 \geq h9$. This allows us to conclude that $h1 \approx h8 \approx h9$ because (i) both $h8$ and $h9$ are either locationally equivalent to or outscoped by $h1$ and (ii) we now know that there is no scopal argument slot (i.e. an argument slot that is filled by a handle rather than an index) outscoped by $h1$. Thus, we have obtained a semantic representation equivalent to the one in (13), whose meaning we have already examined.

5 Constituent coordination in MRS

Constituent coordination can be dealt with in HPSG and MRS by including the following two rules in the grammar. The HEAD feature and the VALENCE feature are features that are allowed in a feature structure that is the value of the CAT feature; the value of the HEAD feature encodes the kinds of grammatical properties that are inherited from a head daughter to its mother, and the value of the VALENCE feature represents valence-related properties. Here and elsewhere, the symbol $\oplus$ is used as an operator that performs list concatenation; for instance, $\langle a \rangle \oplus \langle b, c \rangle = \langle a, b, c \rangle$. The denotation of a variable of the form $x_1 + \cdots + x_n$, where $n > 1$, is assumed to be the sum of the denotations of $x_1, \ldots, x_n$.

(17) The Coordination Rule 1 (preliminary formulation):
Let B_0 be a coordinate structure that consists or n conjuncts or disjuncts, $B_1, \ldots, B_n$

from left to right. Let $c_0, \ldots, c_n$ be the HEAD values of $B_0, \ldots, B_n$ respectively, $v_0, \ldots, v_n$ be the VALENCE values of $B_0, \ldots, B_n$ respectively, $h_1, \ldots, h_n$ be the LTOP values of $B_1, \ldots, B_n$ respectively, h_0 be the SEMHEAD value of B_0, $x_0, \ldots, x_n$ be the INDEX values of $B_0, \ldots, B_n$ respectively, $E_0, \ldots, E_n$ be the EP values of $B_0, \ldots, B_n$ respectively, and $P_0, \ldots, P_n$ be the PHON values of $B_0, \ldots, B_n$ respectively. Then it must be the case that

a. $c_0 = \langle c_1, \ldots, c_n \rangle$,

b. $v_0 = v_1 = \cdots = v_n$,

c. $x_0 = x_1 = \cdots = x_n$,

d. $E_0 = \langle h_0 : C\,(h_1, \ldots, h_n) \rangle \oplus E_1 \oplus \cdots \oplus E_n$, where C is either “and” or “or”, and

e. $P_0 = P_1 \oplus \cdots \oplus P_{n-1} \oplus \langle C \rangle \oplus P_n$.

(18) The Coordination Rule 2 (preliminary formulation):
Let B_0 be a coordinate structure that consists or n conjuncts, $B_1, \ldots, B_n$ from left to right, each of which is a nominal constituent. Let $c_0, \ldots, c_n$ be the HEAD values of $B_0, \ldots, B_n$ respectively, $v_0, \ldots, v_n$ be the VALENCE values of $B_0, \ldots, B_n$ respectively, $h_1, \ldots, h_n$ be the LTOP values of $B_1, \ldots, B_n$ respectively, h_0 be the SEMHEAD value of B_0, $x_0, \ldots, x_n$ be the INDEX values of $B_0, \ldots, B_n$ respectively, $E_0, \ldots, E_n$ be the EP values of $B_0, \ldots, B_n$ respectively, and $P_0, \ldots, P_n$ be the PHON values of $B_0, \ldots, B_n$ respectively. Then it must be the case that

a. $c_0 = \langle c_1, \ldots, c_n \rangle$,

b. $v_0 = v_1 = \cdots = v_n$,

c. $h_0 = h_1 = \cdots = h_n$,

d. $x_0 = x_1 + \cdots + x_n$,

e. $E_0 = E_1 \oplus \cdots \oplus E_n$, and

f. $P_0 = P_1 \oplus \cdots \oplus P_{n-1} \oplus \langle \text{and} \rangle \oplus P_n$.

The Coordinate Structure Rule 1 licenses a coordinate structure expressing logical conjunction or disjunction, and the Coordinate Structure Rule 2 licenses a structure in which multiple nominal expressions are conjoined to form an expression that denotes a group consisting of the denotations of those nominal conjuncts. It needs to be stipulated that the Coordination Rule 1 is an exception to the Semantic Compositionality Principle.

Let us examine some structures that are licensed by these rules. Figure 2 shows one such structure, a local subtree in which conjunction of two Ss is licensed by the Coordination Rule 1. The CAT values associated with the daughter nodes in this figure indicate that the nodes are constituents that are each headed by a verb and are saturated, i.e. are not to be combined with any more complements or subjects; the CAT value of the mother node indicates that the node is a saturated constituent headed by two verbs.[4] The symbol *gtop* stands for a designated handle that is required to be locationally equivalent to the global top handle of the semantic representation as a whole. It is assumed that a proper noun like *Chris* is always associated with an elementary predication that is prefaced with the handle *gtop*. Since $h2 =_q h4$ and $h3 =_q h5$ in Figure 2 because of the LTOP-over-SEMHEAD requirement, the semantic representation associated with the top node is equivalent to (19).

(19) identical(x, Chris) $\wedge$ identical(y, Pat) $\wedge$ and(walks(x), talks(y))

---

[4] See Yatabe (2004) for an explanation of why the CAT value of a coordinate structure needs to be the way it is assumed to be here, although there are some minor differences between the theory depicted here and the theory proposed in that work.

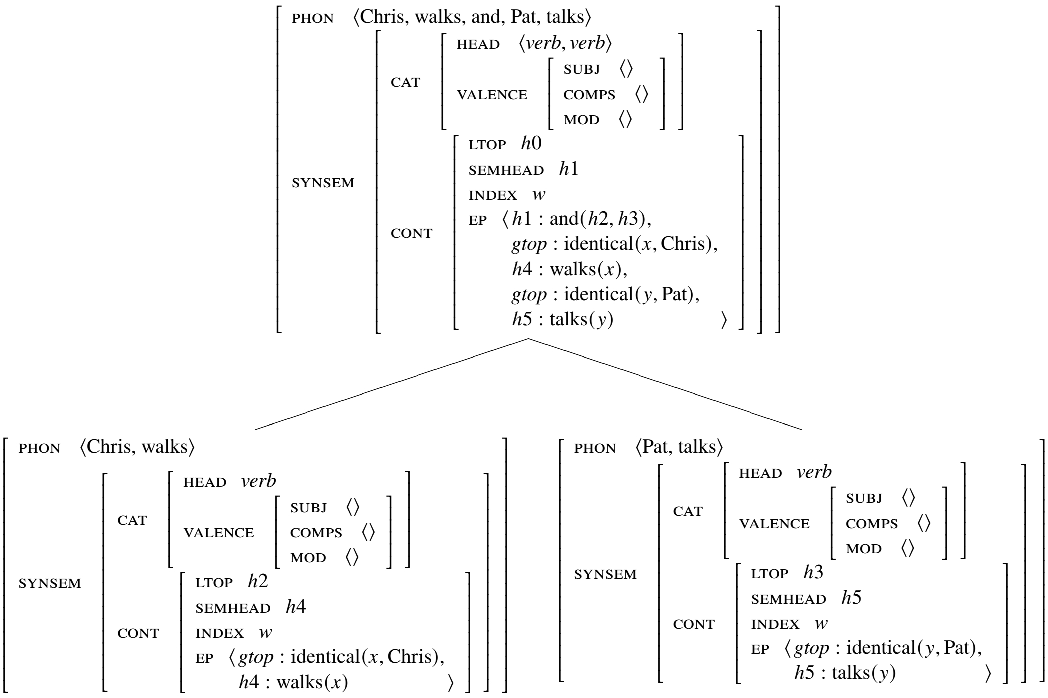

Figure 2: Conjunction of two Ss

I am using the symbol "$\wedge$" to indicate conjunction expressed by shared handles and the symbol "and" to indicate conjunction expressed by an elementary predication whose predicate is *and*.

Figure 3 shows a structure in which two VPs are disjoined by the Coordination Rule 1. The CAT value of the top node indicates that the node is a saturated expression headed by two verbs. The CAT value of the second daughter of the top node indicates that the node is headed by two verbs and is to be combined with a subject but not with any complements. A tag of the form $\boxed{n}$ occurring in multiple places indicates that the same feature structure occurs in those multiple places. The tag $\boxed{1}$ in Figure 3, which is the sole element of the SUBJ list of the second daughter of the top node and which is attached to the SYNSEM value of the first daughter of the top node, means that, in accordance with what is required by the Head-Subject Rule, which is used to license the top local subtree, the sole element of the SUBJ list of the second daughter of the top node is identical to the SYNSEM value of the first daughter of the top node. The local subtree whose root is the second daughter of the top node is the one that is licensed by the Coordination Rule 1 in this figure. Because the rule requires the VALENCE value of the mother to be identical to the VALENCE values of the daughters, the SUBJ values of the daughters are both required to be a list whose sole element is the SYNSEM value of the first daughter of the top node. Since the lexical entry for *talks* requires the INDEX value of the sole element of its SUBJ list to be identical to the argument of the *talks* predicate that appears in its EP list, that argument is forced to be x in this figure. It is because of the same mechanism that the argument of the *walks* predicate also must be x. The LTOP-over-SEMHEAD requirement requires that $h2 =_q h8$, $h3 =_q h9$, and $h5 =_q h7$ (because the LTOP value and the SEMHEAD value of the NP complement of *every*, not shown in the figure, are $h5$ and $h7$ respectively), so the CONT value of the top node in this figure is equivalent to the following, which expresses appropriate truth conditions.

(20) every(x, child(x), or(talks(x), walks(x)))

Notice that there is virtually no difference between the way coordination of Ss is treated and the way coordination of VPs is treated. In both cases, the coordinate structure introduces

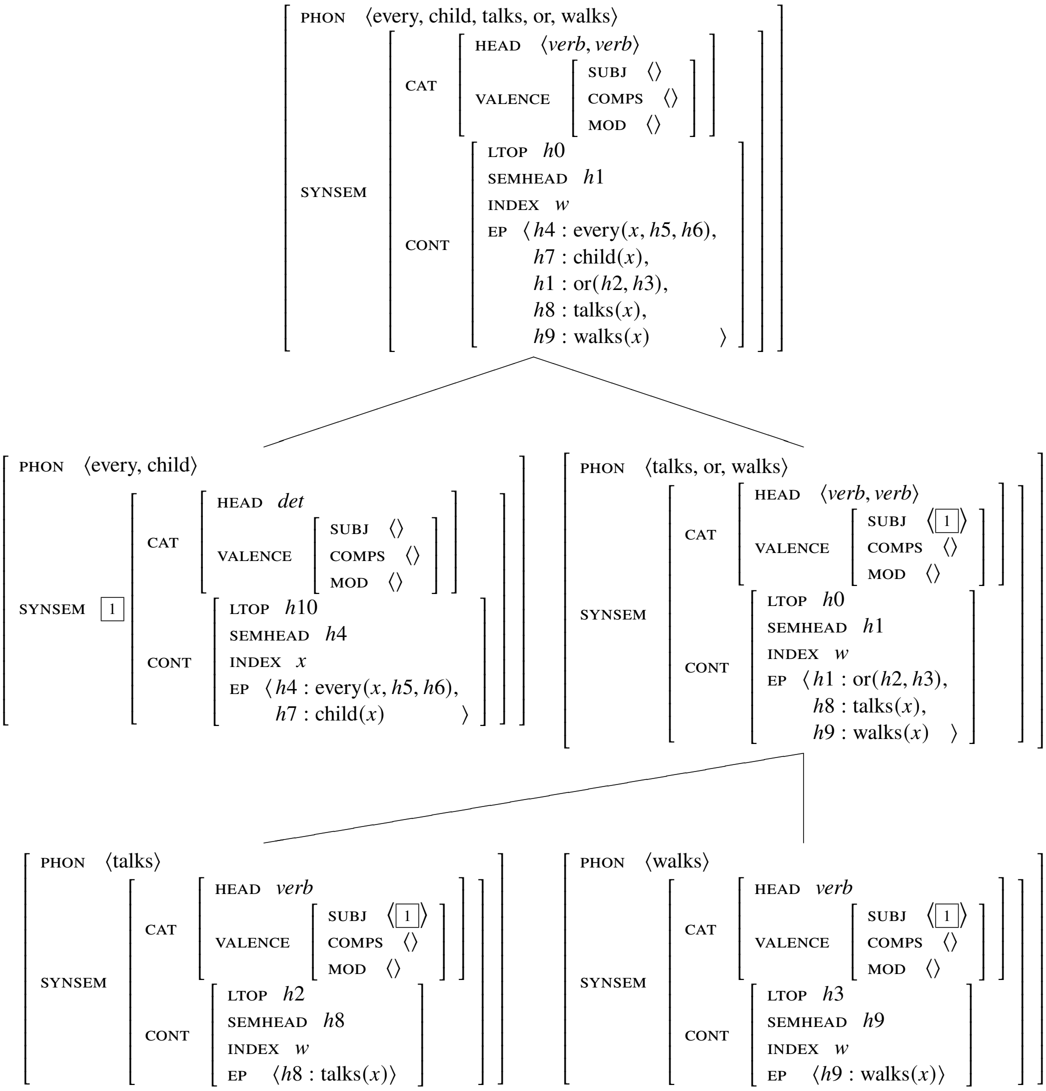

Figure 3: Disjunction of two VPs

into the semantic representation an elementary predication whose arguments are the LTOP values of the constituents being coordinated.

Figure 4 shows a structure in which conjunction of two proper nouns is licensed by the Coordination Rule 2. The local subtree whose root is the first daughter of the top node is the one that is licensed by the Coordination Rule 2. In this local subtree, two proper nouns whose indices are x and y respectively are conjoined to produce a DP pronounced *Chris and Pat*, whose index is $x + y$. The CONT value associated with the top node in the figure is equivalent to (21).

(21) identical(x, Chris) $\wedge$ identical(y, Pat) $\wedge$ talk($x + y$)

Figure 5 shows the structure of the sentence *Ten men and women talk*, in which conjunction of two common nouns is licensed by the Coordination Rule 2. I assume here that the word *ten* in the sentence is an adjective and that there is an unpronounced determiner immediately preceding that adjective. The local subtree at the very bottom of the figure is the one licensed by the Coordination Rule 2. In that local subtree, the noun *men*, whose index is x, is conjoined with the noun *women*, whose index is y, to produce an NP pronounced *men and women*, whose index is $x + y$. The CONT value of the top node in the figure is equivalent to (22), because the LTOP-over-SEMHEAD requirement implies that $h6$ must either

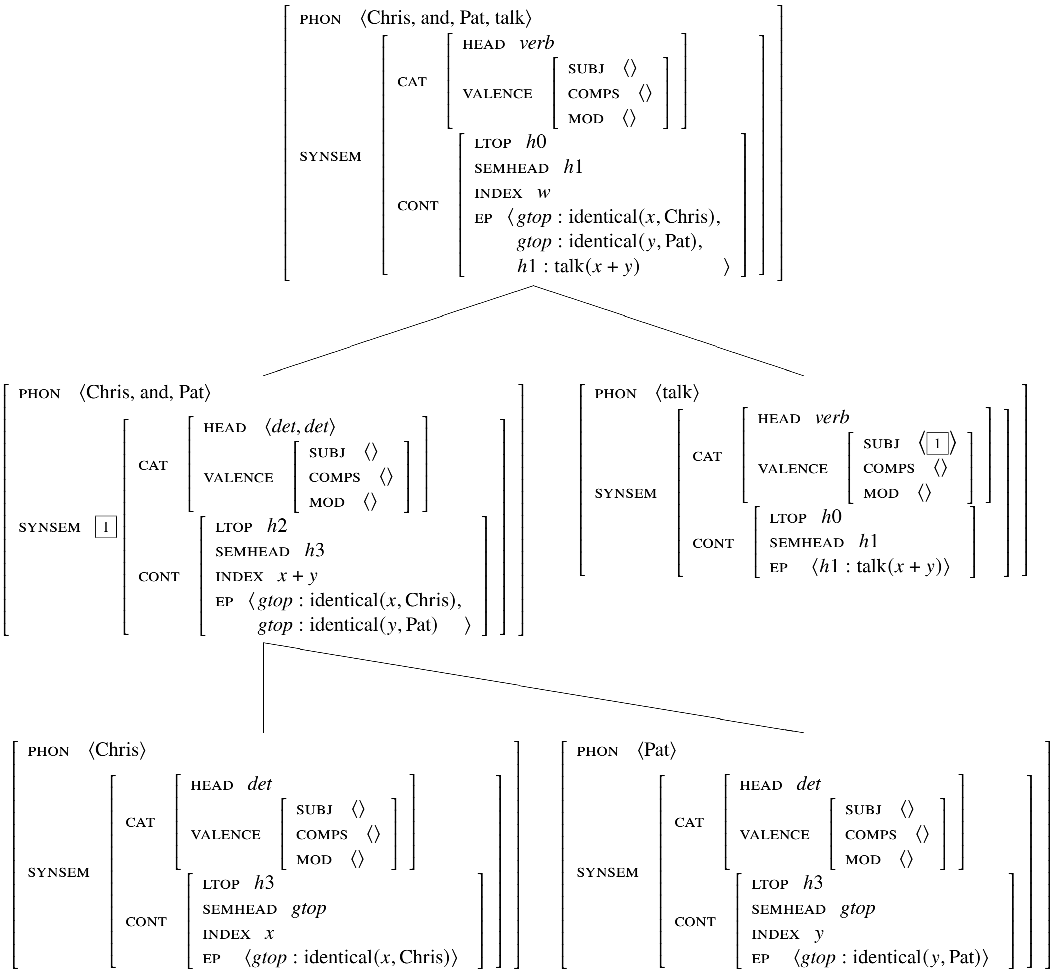

Figure 4: Group-forming conjunction of two proper nouns

be locationally equivalent to or outscope $h2$, $h3$, and $h4$.

(22) $\text{some}(x + y, \text{ten}(x + y) \wedge \text{men}(x) \wedge \text{women}(y), \text{talk}(x + y))$

Notice that there is almost no difference between the way group-forming conjunction of proper names is treated and the way group-forming conjunction of common nouns is treated. In both cases, conjunction produces a nominal expression whose index is of the form $x_1 + \cdots + x_n$, where x_1, ..., x_n are the indices of the nominal constituents being coordinated.

In the proposed theory, it would be surprising if there were a language in which different morphemes were used to express coordination of intransitive verbs and coordination of sentences or to express group-forming conjunction of proper names and group-forming conjunction of common nouns.

6 Coordination of scopal predicates

The MRS-based account of coordination presented in the previous section in fact has one seriously flaw; it is incapable of dealing with some instances of coordination of scopal predicates, that is, predicates which take one or more handles as arguments. More specifically, the theory cannot deal with a situation where two scopal predicates share the same handle as their arguments, because by assumption a set of elementary predications cannot be used to fill more than one argument slot. For instance, a sentence like (23) cannot be associated with

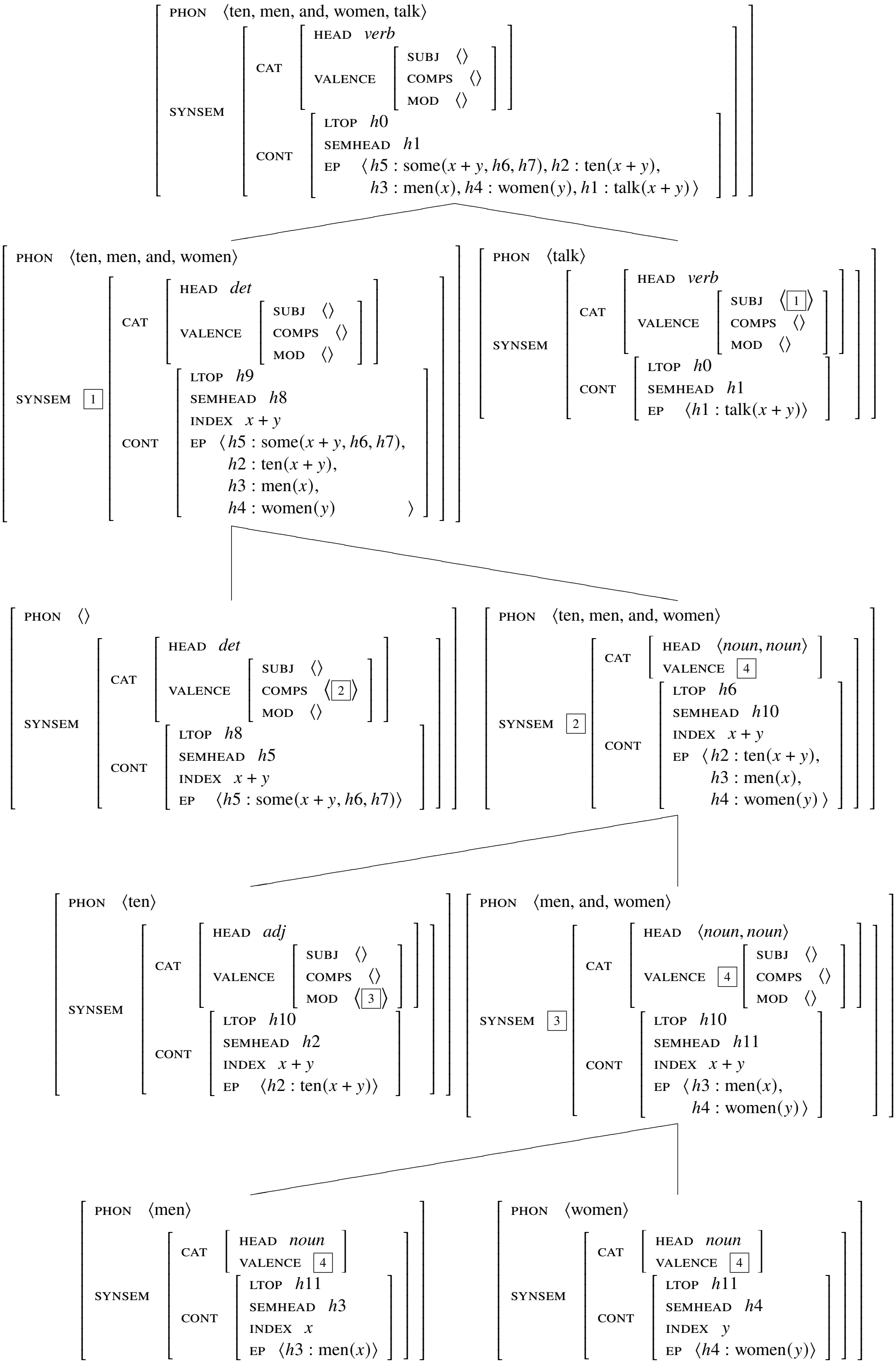

Figure 5: Group-forming conjunction of two common nouns

a semantic representation like (24), because in (24) the elementary predication "snore(x)" is used to fill two argument slots.

(23) Many women and most men snore.

(24) and(many(x, women(x), snore(x)), most(x, men(x), snore(x)))

The theory is likewise unable to assign to a sentence like (25) a semantic representation like (26), in which the elementary predication "snores(y)" is used to fill two argument slots.

(25) Chris suspects and fears that Pat snores.

(26) identical(x, Chris) $\wedge$ identical(y, Pat)
$\wedge$ and(suspects(x, snores(y)), fears(x, snores(y)))

Three modifications need to be made to the theory in order to deal with this issue. First, we need to drop the assumption that an elementary predication cannot be used to fill two or more argument slots at the same time. I propose to replace that assumption with the constraint stated in (27).

(27) Let R be the root node of a phrase-structure tree, T. Then no two scopal argument slots contained in the EP list associated with R are allowed to be materially identical unless they are required to be nominally identical in T.

(28) We say that two scopal argument slots S_1 and S_2 are *materially identical* if and only if the handle representing S_1 and the handle representing S_2 are locationally equivalent to each other.

(29) We say that two scopal argument slots S_1 and S_2 are *required to be nominally identical in* a phrase-structure tree T if and only if
  (i) S_1 and S_2 are represented by the same handle, h, and
  (ii) there is no way to replace one or more occurrences of h in T with a different handle h' such that (a) the replacement either changes the handle representing S_1 from h to h' while leaving the handle representing S_2 unchanged or changes the handle representing S_2 from h to h' while leaving the handle representing S_1 unchanged and (b) the phrase-structure tree that results from the replacement is grammatical.

A semantic representation like (26) becomes possible as a result of this first modification, because in the phrase-structure tree representing the structure of (25), the second (i.e. the non-experiencer) argument slot of *suspects* and the second (i.e. the non-experiencer) argument slot of *fears* are required to be nominally identical (primarily due to the fact that the Coordination Rule 1 demands that the VALENCE values of conjuncts be identical to each other) and are therefore allowed to be materially identical.

Second, we need to add to the grammar a mechanism that requires two or more quantifiers that are coordinated with each other to have the same scope. In order to achieve this goal, I propose to add a new feature, SCOPE, to feature structures that serve as values of the VALENCE feature. In addition, let us assume (i) that the lexical entry of each quantificational nominal (such as the determiner *every*) requires the value of this SCOPE feature to be identical to the scope argument of the elementary predication representing the quantificational meaning (which is the third argument in the case of an elementary predication whose predicate is *every*) and (ii) that every other lexical entry requires the value of its own SCOPE feature to be "none". Given this setup, quantifiers that are coordinated are always required to have the same scope because the scope of each quantifier is now represented in the VALENCE value of that quantifier and the Coordination Rule 1 requires the constituents being coordinated to share the same VALENCE value. Consequently, a semantic representation like (24) can

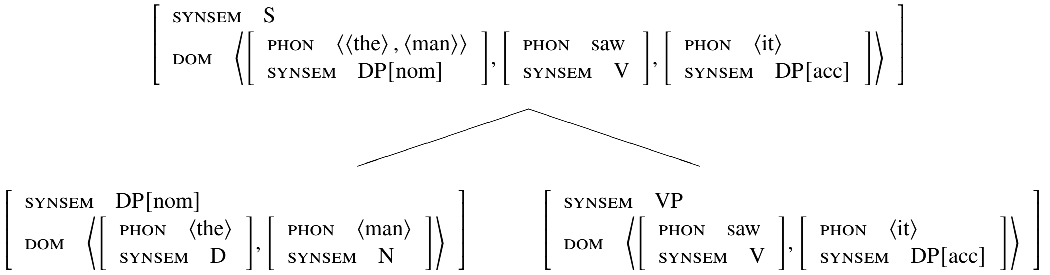

Figure 6: Total compaction of a DP and liberation of a VP

now be associated with sentence (23) because the third argument slot of *many* and the third argument slot of *most* are required to be nominally identical in the phrase-structure tree representing the structure of (23) and are thus allowed to be materially identical.

And third, the variable-binding condition must be relaxed so that two quantificational elementary predications are allowed to share the same VARIABLE value when they share the same SCOPE value.

7 Linearization-based HPSG

At this point, in order to be able to deal with linguistic phenomena involving discontinuous constituency such as right-node raising and left-node raising, we need to add to our theory some mechanisms that have been developed in a theoretical framework called Linearization-based HPSG, the version of HPSG proposed in works such as Reape (1994) and Kathol and Pollard (1995). In this framework, each node in a syntactic phrase-structure tree is associated with what is called an *order domain*, which can be viewed as a list of prosodic constituents that the words dominated by that node are assembled into. An order domain is a list of *domain objects*, and is given as the value of the DOM feature. Departing from an assumption adopted in earlier versions of Linearization-based HPSG, I will assume that morphosyntactic constituency is represented not by an unordered tree but by a tree, and that the order of domain objects in an order domain largely reflects the order of the nodes that have given rise to those domain objects (see Yatabe and Tam (2021, Appendix A)).

I will use some concrete examples to explicate how the content of the order domain associated with each node is determined. Figure 6 shows part of the structure assigned to the English sentence *The man saw it.* What is shown is the local subtree in which the subject DP and the VP combine to become the matrix clause. The order domain (i.e. the DOM value) of the VP node consists of two domain objects, one that is pronounced *saw*, and the other one that is pronounced *it*; this indicates that this VP is to be pronounced *saw it.* Likewise, the order domain of the DP node indicates that this DP is to be pronounced *the man*, and the order domain of the S node indicates that the S node is to be pronounced *The man saw it.*

Notice that the subject DP node, whose order domain contains two domain objects, contributes to the order domain of the S node only one domain object, which is pronounced *the man.* What is at work here is an operation called *total compaction.* (30) is a simplified description of the way the total compaction operation takes a feature structure associated with a node and turns it into a single domain object.

(30) Total compaction (Preliminary description)

$$\begin{bmatrix} \text{SYNSEM} & \alpha_0 \\ \text{DOM} & \left\langle \begin{bmatrix} \text{PHON} & \beta_1 \\ \text{SYNSEM} & \alpha_1 \end{bmatrix}, \cdots, \begin{bmatrix} \text{PHON} & \beta_n \\ \text{SYNSEM} & \alpha_n \end{bmatrix} \right\rangle \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \text{PHON} & \langle \beta_1, \cdots, \beta_n \rangle \\ \text{SYNSEM} & \alpha_0 \end{bmatrix}$$

What is shown on the left of the arrow is the input to the operation; the input is a feature structure associated with a node. On the right of the arrow is shown the output of the operation; the output is a domain object. The domain object that is created by totally compacting (a feature structure associated with) a node X is placed in the order domain of the mother of X. In Figure 6, the domain object that is created by totally compacting the subject DP has been placed in the order domain of the matrix S.

We say that a node has been *liberated* when the node is not totally compacted and all the domain objects in the order domain of the node are inherited unaltered by the order domain of the mother of that node. The VP in Figure 6 is liberated. The two domain objects in the order domain of the VP node are both integrated, unaltered, into the order domain of the S node.

There is a third process that a node may undergo, besides total compaction and liberation. We say that (a feature structure associated with) a node has been *partially compacted* when (i) zero or more domain objects are excised from its order domain and (ii) the feature structure thus altered is compacted. Note that, by definition, total compaction is in fact a type of partial compaction. When a node is partially compacted, the domain objects that were excised from its order domain (if any) are inherited by the order domain of the mother of that node, and the domain object that is newly created by compaction is also placed in the order domain of the mother. We will say that those domain objects that were excised from the order domain of a node and inherited by the order domain of the mother of that node have *escaped compaction*, and we will say that the domain objects that were not excised and thus were part of the feature structure that was compacted (in stage (ii) of partial compaction) have *undergone compaction*.

In the original definition of partial compaction presented in Kathol and Pollard (1995), it is assumed that only domain objects corresponding to extraposable types of expressions (such as domain objects corresponding to relative clauses in English) can escape compaction, but here we drop that restriction and assume that, in a head-first language like English, any domain object can escape compaction as long as it is not the leftmost element of an order domain. (31) describes the way the partial compaction operation takes the feature structure associated with a node and turns it into one or more domain objects, which are to be placed in the order domain of the mother of that node.

(31) Partial compaction for head-first languages (Preliminary description):

$$\begin{bmatrix} \textsc{synsem} & \alpha_0 \\ \textsc{dom} & \left\langle \begin{bmatrix} \textsc{phon} & \beta_1 \\ \textsc{synsem} & \alpha_1 \end{bmatrix}, \cdots, \begin{bmatrix} \textsc{phon} & \beta_n \\ \textsc{synsem} & \alpha_n \end{bmatrix} \right\rangle \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \textsc{phon} & \langle \beta_1, \cdots, \beta_i \rangle \\ \textsc{synsem} & \alpha_0 \end{bmatrix}, \begin{bmatrix} \textsc{phon} & \beta_{i+1} \\ \textsc{synsem} & \alpha_{i+1} \end{bmatrix}, \cdots, \begin{bmatrix} \textsc{phon} & \beta_n \\ \textsc{synsem} & \alpha_n \end{bmatrix}$$

What is shown on the left of the arrow is the input to the partial compaction operation, i.e. the feature structure associated with a node, whose DOM value contains n domain objects; shown on the right of the arrow are domain objects that will be placed in the order domain of the mother of that node. The first domain object on the right of the arrow is the result of compacting the feature structure obtained by excising the rightmost $n - i$ domain objects from the order domain of the input, and the remaining domain objects on the right of the arrow are those $n - i$ domain objects, i.e. the ones that have escaped compaction.

Various types of extraposition constructions may result when an expression is partially and not totally compacted and surfaces as a discontinuous constituent. Figure 7 shows how the English extraposition construction can be generated via partial compaction. What is shown in the figure is a local subtree in which a subject DP *a man who was wearing a black cloak* and a VP *entered* combine to become a sentence *A man entered who was wearing a black cloak*. Here, the subject DP has been partially compacted; the relative clause has

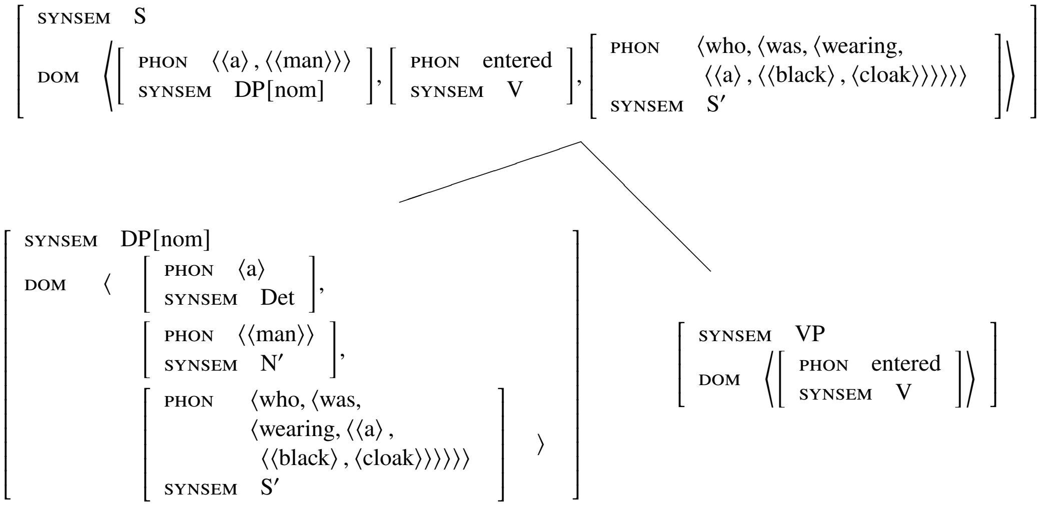

Figure 7: Partial compaction of a DP

escaped compaction and appears in the sentence-final position. As we will see shortly, something resembling partial compaction plays a crucial role in the proposed account of right- and left-node raising as well.

What happens to each node must conform to the set of constraints shown in (32).[5]

(32) a. In a head-complement structure whose head is not nominal and in a head-subject structure, the head is liberated and the non-head is partially compacted.

b. In a head-adjunct structure whose head is not nominal, the head and the adjunct are both partially compacted.

c. In a headed structure whose head is a nominal expression (i.e. an N, a D, or a projection of either an N or a D) and whose non-head is not a marker, the head is totally compacted and the non-head is partially compacted.

8 Compaction as the driver of meaning assembly

The framework of linearization-based HPSG in itself does not necessitate any modification to the way semantic composition is performed, but in order to allow this theoretical framework to achieve more comprehensive descriptive adequacy than competing frameworks, we do need to modify slightly the way MRS is integrated into Linearization-based HPSG.

I will start by examining how semantic composition is carried out in Linearization-based HPSG as it has been presented so far. Figure 8 shows the syntactic phrase-structure tree assigned to the sentence *Some boy saw every girl* in this theory, a structure in which a determiner *every* and a noun *girl* combine to form a DP, which combines with a transitive verb *saw* to become a VP, which then combines with a subject DP *some boy* (whose internal structure is suppressed in this figure) to become a sentence. Here and in later illustrations, elementary predications are represented as feature structures just like other objects used to describe grammatical structure. A feature structure representing an elementary predication has the feature RELN (which is short for *relation*), whose value represents the predicate involved, the feature HNDL (which is short for *handle*), whose value represents the handle attached to the elementary predication, and other features whose values represent the arguments contained in the elementary predication.

---

[5] In these constraints, nominal expressions and non-nominal headed expressions are treated differently. The motivation for this distinction has to do with a certain restriction on long-distance scrambling in Japanese, as explained in Yatabe (2009, fn. 6). The distinction plays no role in this article.

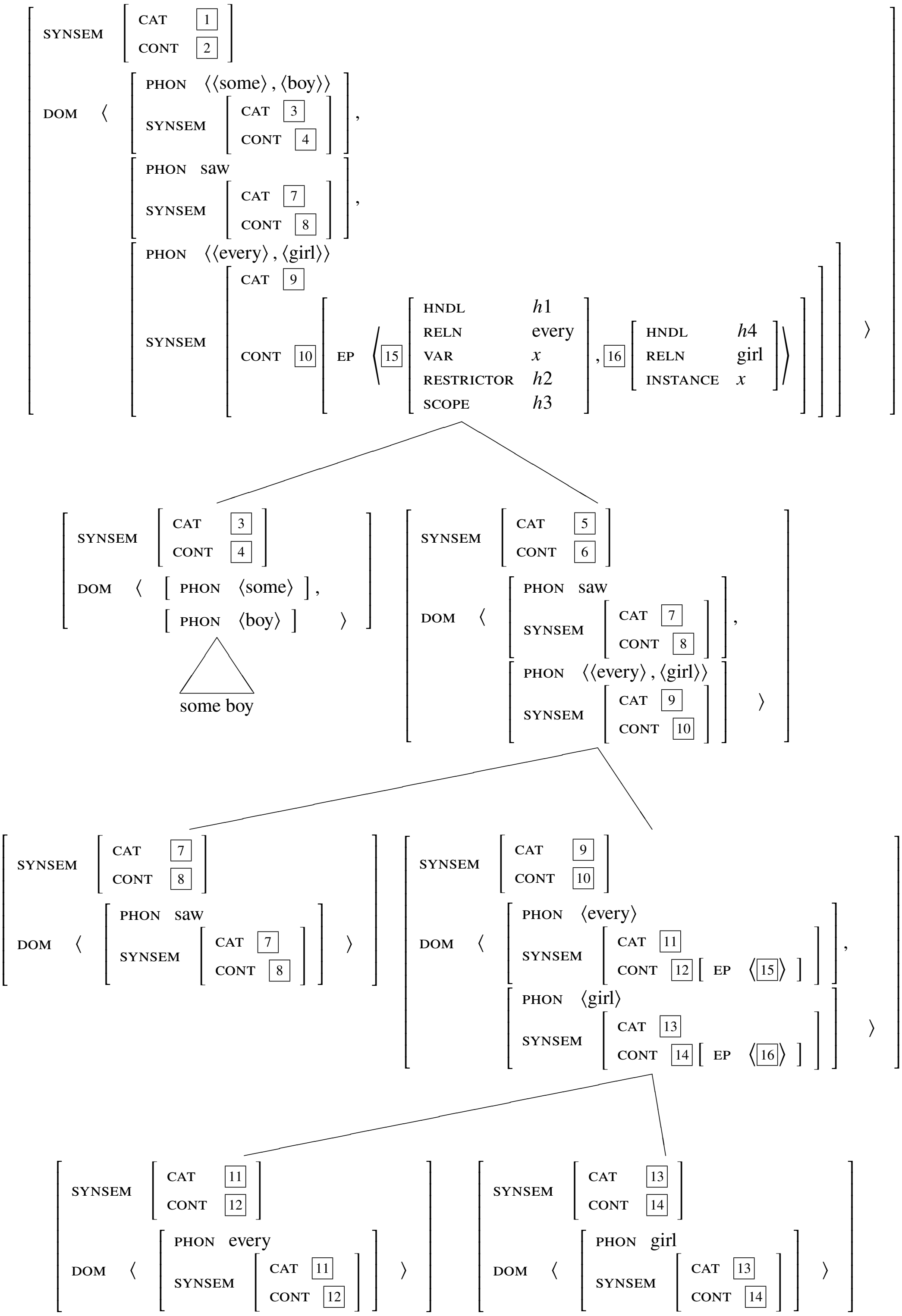

Figure 8: A schematic representation of the structure assigned to the sentence *Some boy saw every girl* in the standard versions of Linearization-based HPSG

What is to be noted here is that, in the theory as it has been presented so far, meaning assembly proceeds along two routes, so to speak. On the one hand, the meaning of each non-leaf node is computed on the basis of the meaning of its daughters. But at the same time, as a side effect of this, the meaning of successively larger domain objects is computed as well, since the CONT value of each domain object mirrors the CONT value of the node whose compaction gave rise to that domain object. For instance, let us see how the meaning of *every* and that of *girl* are combined to yield the meaning of the phrase *every girl* in Figure 8. This process takes place in the local subtree at the very bottom, where the meaning of the DP node *every girl* (i.e. the denotation of $\boxed{10}$) is computed on the basis of the meaning of the D node *every* (i.e. the denotation of $\boxed{12}$) and the meaning of the N node *girl* (i.e. the denotation of $\boxed{14}$). (The denotation of $\boxed{10}$ is schematically shown within the top node, and the denotations of $\boxed{12}$ and $\boxed{14}$ are schematically shown within the object DP node.) But at the same time, if you ignore the SYNSEM|CONT values of nodes and instead focus on the CONT values of domain objects in the figure, it looks as though the semantic content of the second domain object in the order domain of the VP node (namely the domain object to be pronounced "every girl") had been computed by combining the semantic content of the first and the second domain object in the order domain of the object DP node (namely the two domain objects to be pronounced "every" and "girl" respectively.)

We will eliminate this duplication in meaning assembly by assuming that the EP values associated with nodes are always empty lists while continuing to use the EP values associated with domain objects as loci for representing the semantic interpretation assigned to those objects. More specifically, we will (i) excise the Semantic Compositionality Principle from the theory and (ii) redefine the compaction operation as in Figure 9.[6]

Before explaining the content of Figure 9, let me introduce the notion of *sort* at this point. Feature structures that are used in the grammar are categorized into sorts. A feature structure that serves as a domain object belongs to the sort *dom-obj*. A feature structure that is associated with a node in a phrase-structure tree belongs to the sort *sign*, and is accordingly referred to as a sign.

What Figure 9 means is that, when a sign that has the form described in the first line is compacted, the domain object that is produced as a result must have the form described in the second line. The definition is a specification of what happens to a sign when it is totally compacted as well as what happens to a sign in the second stage of partial compaction (the first stage being excision of some domain objects from the order domain of the sign); as such, it is meant to replace the simplified description of total compaction shown in (30) as well as part of the description of partial compaction given in (31) above. Each time a sign is compacted according to this definition, a new domain object is created whose EP value is the result of concatenating the EP value of each domain object in the order domain of that sign. The function f in the definition is a function that constructs an appropriate phonological representation out of smaller phonological representations. For the sake of concreteness, let us assume the following, which is in fact adequate in most cases.

(33) $$f\left(\boxed{d_1}, \cdots, \boxed{d_n}\right) = \left\langle \boxed{d_1}, \cdots, \boxed{d_n} \right\rangle$$

Figures 10, 11, and 12 show part of the lexical entries that I assume for the words *saw*, *every*, and *girl*, respectively. In each lexical entry, the SEMHEAD value of the sign is identified with the HNDL value of an elementary predication contained in the EP value of a domain object. The SEMHEAD value of a leaf node is thus always linked to the HNDL value

---

[6] The definition of compaction given in the present document is simpler than the definitions provided in Yatabe (2001) and Yatabe and Tam (2021) for the following three reasons. First, we do not need to have the H-CONS feature now that we have the LTOP-over-SEMHEAD requirement. Second, I am now assuming that the KEY feature is a HEAD feature, not part of the CONTENT value. And third, unlike the earlier formulations, the definition given in Figure 9 does not encode the view that there are syntactic constituents that function as scope islands.

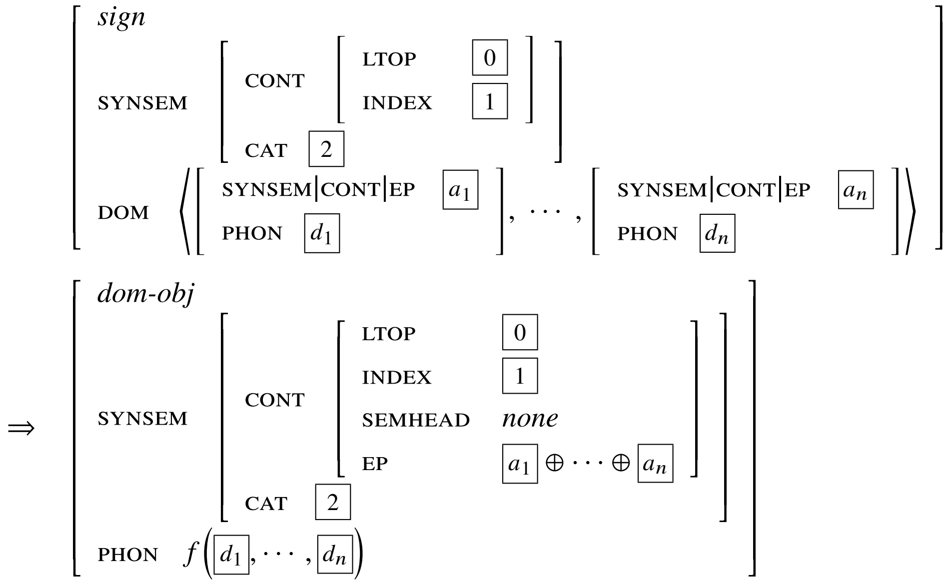

Figure 9: Definition of compaction

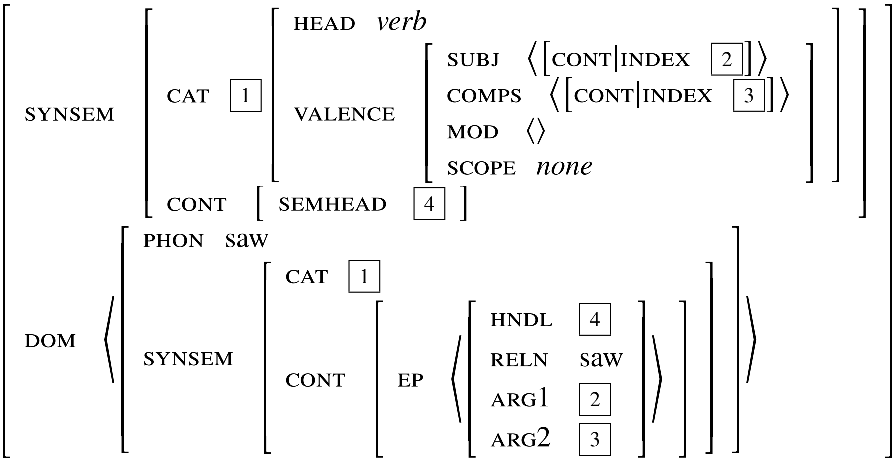

Figure 10: Part of the lexical entry for *saw*

of some elementary predication. In the lexical entry for *saw*, the indices of the subject and the object are identified respectively with the first argument and the second argument of the *saw* relation in the EP value of the domain object. In the lexical entry for *every*, the LTOP value and the INDEX value of the complement nominal are identified respectively with the RESTRICTOR argument and the VAR argument of the *every* relation in the EP value of the domain object; the VAR argument represents the variable bound by the quantifier.

The semantic content of a sentence as a whole is obtained by first totally compacting the top node of the syntactic representation and then resolving the resulting MRS representation in conformity with all the grammatical constraints including the LTOP-over-SEMHEAD requirement.

The Coordination Rules 1 and 2 need to be modified slightly at this point in order to be consistent with the rest of the theory.

(34) The Coordination Rule 1:
Let B_0 be a coordinate structure that consists or n conjuncts or disjuncts, $B_1, \ldots, B_n$ from left to right. Let $c_0, \ldots, c_n$ be the HEAD values of $B_0, \ldots, B_n$ respectively, $v_0, \ldots, v_n$ be the VALENCE values of $B_0, \ldots, B_n$ respectively, $h_1, \ldots, h_n$ be the LTOP values of $B_1, \ldots, B_n$ respectively, h_0 be the SEMHEAD value of B_0, and $x_0, \ldots, x_n$ be

$$\begin{bmatrix} \textsc{synsem} & \begin{bmatrix} \textsc{cat} & \boxed{1}\begin{bmatrix} \textsc{head} & \textit{det} \\ \textsc{valence} & \begin{bmatrix} \textsc{subj} & \langle\rangle \\ \textsc{comps} & \left\langle \begin{bmatrix} \textsc{cont} & \begin{bmatrix} \textsc{ltop} & \boxed{4} \\ \textsc{index} & \boxed{3} \end{bmatrix} \end{bmatrix} \right\rangle \\ \textsc{mod} & \langle\rangle \\ \textsc{scope} & \boxed{5} \end{bmatrix} \end{bmatrix} \\ \textsc{cont} & \begin{bmatrix} \textsc{semhead} & \boxed{2} \\ \textsc{index} & \boxed{3} \end{bmatrix} \end{bmatrix} \\ \textsc{dom} & \left\langle \begin{bmatrix} \textsc{phon} & \text{every} \\ \textsc{synsem} & \begin{bmatrix} \textsc{cat} & \boxed{1} \\ \textsc{cont} & \begin{bmatrix} \textsc{ep} & \left\langle \begin{bmatrix} \textsc{hndl} & \boxed{2} \\ \textsc{reln} & \text{every} \\ \textsc{var} & \boxed{3} \\ \textsc{restrictor} & \boxed{4} \\ \textsc{scope} & \boxed{5} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

Figure 11: Part of the lexical entry for *every*

$$\begin{bmatrix} \textsc{synsem} & \begin{bmatrix} \textsc{cat} & \boxed{1}\begin{bmatrix} \textsc{head} & \textit{noun} \\ \textsc{valence} & \begin{bmatrix} \textsc{subj} & \langle\rangle \\ \textsc{comps} & \langle\rangle \\ \textsc{mod} & \langle\rangle \\ \textsc{scope} & \textit{none} \end{bmatrix} \end{bmatrix} \\ \textsc{cont} & \begin{bmatrix} \textsc{semhead} & \boxed{2} \\ \textsc{index} & \boxed{3} \end{bmatrix} \end{bmatrix} \\ \textsc{dom} & \left\langle \begin{bmatrix} \textsc{phon} & \text{girl} \\ \textsc{synsem} & \begin{bmatrix} \textsc{cat} & \boxed{1} \\ \textsc{cont} & \begin{bmatrix} \textsc{ep} & \left\langle \begin{bmatrix} \textsc{hndl} & \boxed{2} \\ \textsc{reln} & \text{girl} \\ \textsc{instance} & \boxed{3} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

Figure 12: Part of the lexical entry for *girl*

the INDEX values of $B_0, \dots, B_n$ respectively, Then it must be the case that

a. $c_0 = \langle c_1, \dots, c_n \rangle$,

b. $v_0 = v_1 = \cdots = v_n$,

c. $x_0 = x_1 = \cdots = x_n$, and

d. the DOM value of B_0 is

$$\left\langle \begin{bmatrix} \textsc{phon} & \langle\rangle \\ \textsc{synsem}|\textsc{cont} & \begin{bmatrix} \textsc{ep} & \left\langle \begin{bmatrix} \textsc{hndl} & h_0 \\ \textsc{reln} & C \\ \textsc{coordinated} & \langle h_1, \dots, h_n \rangle \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \right\rangle \oplus \langle B'_1, \dots, B'_n \rangle,$$

where C is either “and” or “or”, B'_1, …, B'_{n-1} are the domain objects obtained by compacting $B_1, \dots, B_{n-1}$ respectively, and B'_n is the domain object obtained by first compacting B_n to create a new domain object, B''_n, and then adding C at the beginning of the PHON value of B''_n.

(35) The Coordination Rule 2:

Let B_0 be a coordinate structure that consists or n conjuncts, $B_1, \dots, B_n$ from left to right, each of which is a nominal constituent. Let $c_0, \dots, c_n$ be the HEAD values of $B_0, \dots, B_n$ respectively, $v_0, \dots, v_n$ be the VALENCE values of $B_0, \dots, B_n$

respectively, $h_1, \ldots, h_n$ be the LTOP values of $B_1, \ldots, B_n$ respectively, h_0 be the SEMHEAD value of B_0, and $x_0, \ldots, x_n$ be the INDEX values of $B_0, \ldots, B_n$ respectively. Then it must be the case that

a. $c_0 = \langle c_1, \ldots, c_n \rangle$,

b. $v_0 = v_1 = \cdots = v_n$,

c. $h_0 = h_1 = \cdots = h_n$,

d. $x_0 = x_1 + \cdots + x_n$, and

e. the DOM value of B_0 is $\langle B'_1, \ldots, B'_n \rangle$, where $B'_1, \ldots, B'_{n-1}$ are the domain objects obtained by compacting $B_1, \ldots, B_{n-1}$ respectively and B'_n is the domain object obtained by first compacting B_n to create a new domain object, B''_n, and then adding "and" at the beginning of the PHON value of B''_n.

9 Right-node raising

The modifications that I made to the theory in the previous section do not affect the predictions made by the theory as long as we are dealing only with simple grammatical phenomena; all that has changed is the route through which elementary predications are passed up from the leaf nodes to the root node. The effect of the modifications becomes crucial, however, when we deal with phenomena such as right-node raising and left-node raising.

Following the established tradition, I refer to the operation that fuses the right edges of phrases (typically conjuncts) to produce a sentence like *Keats traveled to, and Chapman returned from, the Isle of Capri* as right-node raising (RNR), and to the operation that fuses the left edges of phrases (typically conjuncts) to produce a sentence like *John drove to Chicago in the morning and Detroit in the afternoon* as left-node raising (LNR). The theory of RNR and LNR presented in Yatabe (2001) and modified in later works such as Yatabe (2012) posits that there are several types of RNR and LNR, some of which are meaning-preserving and some of which are potentially meaning-changing. Here I will focus on those types of RNR that are potentially meaning-changing, exemplified by the following sentences.

(36) Some woman likes and some man detests every saxophonist. (from Steedman (2000, p. 78))

(37) I talked to, without actually meeting, all of the members who voted against Hinkly. (from Williams (1990))

(38) John avoided and Bill ignored $\left\{ \begin{array}{c} \text{similar issues} \\ \text{the same man} \end{array} \right\}$. (from Jackendoff (1977, p. 192))

(39) John defeated, whereas Mary lost to, the exact same opponent. (from Kubota and Levine (2020, p. 123))

In this theory, both meaning-preserving types and potentially meaning-changing types of RNR and LNR are dealt with in a way similar to the way extraposition from NP is dealt with. Expressions that involve RNR or LNR are assumed to have the same constituent structure as the corresponding expressions that do not involve RNR or LNR; the former and the latter are taken to differ from each other only in their order domains. For instance, sentence (36) is analyzed as in Figure 13, when a potentially meaning-changing type of RNR is assumed to be implicated in licensing it. The sentence is given the same constituent structure as the sentence *Some woman likes every saxophonist and some man detests every saxophonist*, but some of the nodes in the sentence are given order domains that are different from the order domains of the corresponding nodes in the RNR-free variant of the sentence. What is shown in Figure 13 is the local subtree in which two Ss are conjoined and the object DP

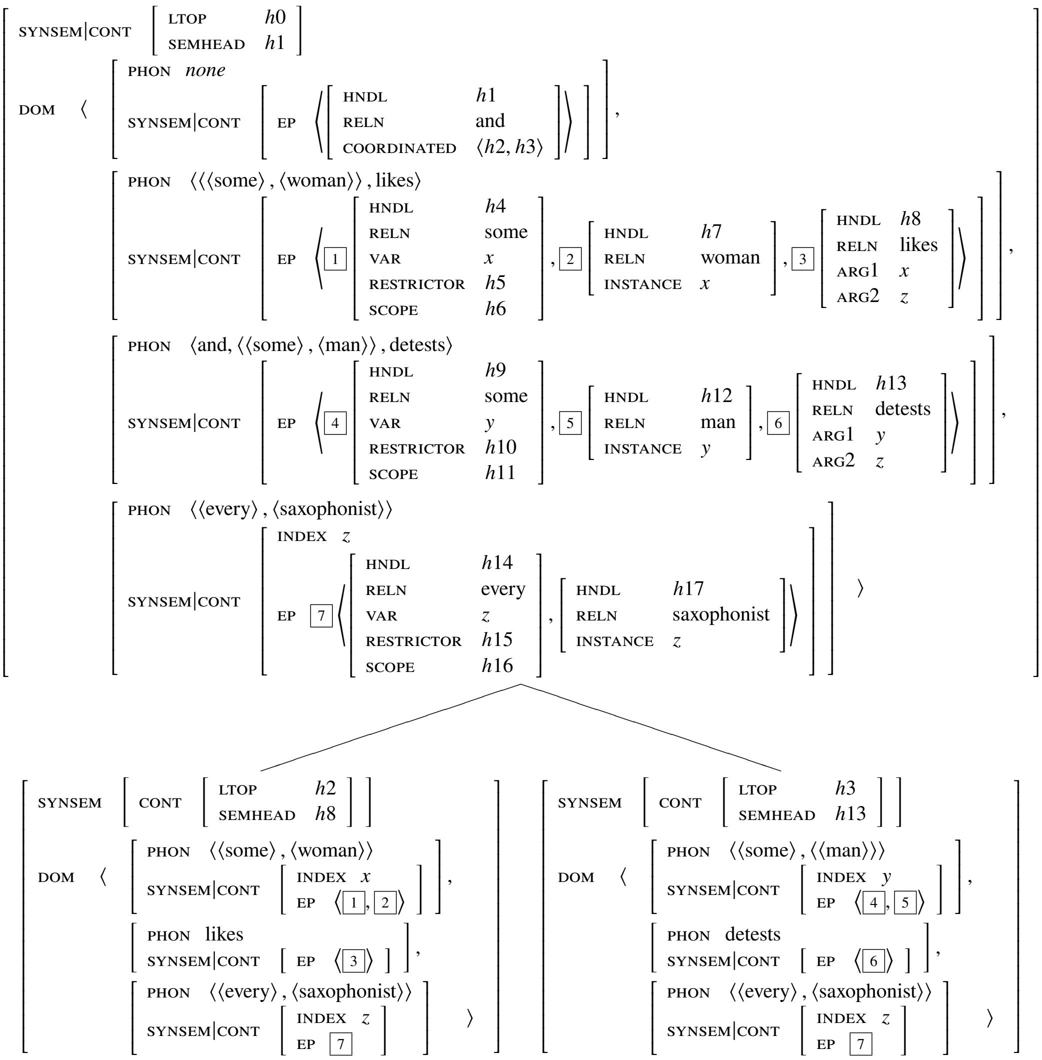

Figure 13: Part of a structure that can be assigned to example (36)

is right-node-raised. The right-node-raised material escapes compaction and continues to be an independent domain object in the order domain of the mother node while the rest of each conjunct is compacted and becomes a single domain object. Thus, the order domain (i.e. the DOM value) of the mother node in the figure ends up consisting of the following four domain objects: (i) a phonologically empty domain object that carries the meaning of conjunction, introduced into the representation by the Coordination Rule 1, (ii) the domain object obtained by compacting the part of the first conjunct that is not right-node-raised, (iii) the domain object obtained by compacting the part of the second conjunct that is not right-node-raised, and (iv) the domain object that is right-node-raised.

In order to license a structure like that shown in Figure 13, it is assumed that the constraints stated in (32) and what the Coordination Rules 1 and 2 say about order domains can optionally be overridden by the following constraint. (See Yatabe (2012) for a more complete formulation that takes into account not just potentially meaning-changing RNR but also LNR and meaning-preserving RNR.)

(40) Let B_0 be a node that consists or n daughters, $B_1, \ldots, B_n$ from left to right. Then the DOM value of B_0 can be $\langle d_1, \ldots, d_n, r'_1, \ldots, r'_m \rangle$, where

  (a) for each i such that $1 \leq i \leq n - 1$, d_i is the domain object obtained by (i) creating a variant of B_i in which m domain objects, $r^i_1, \ldots, r^i_m$ from left to right, have been excised from the right edge of the DOM value of B_i and (ii) compacting the resulting variant of B_i,

(b) d_n is the domain object obtained by first applying the procedure stated in (a) above to B_n and then, if B_0 is a coordinate structure, adding the appropriate conjunction word form (i.e. either "and" or "or") at the beginning of the PHON value of the resulting domain object, and

(c) for each i such that $1 \leq i \leq m$, r'_i is the domain object obtained by fusing n domain objects $r_i^1, \ldots, r_i^n$.

The term *fusing*, which is used in (40), is defined roughly as follows. (See Yatabe (2012) and Yatabe and Tam (2021, fn. 13) for a more complete definition.)

(41) Let $d_0, \ldots, d_n$ be domain objects, $p_0, \ldots, p_n$ be their respective PHON values, $x_0, \ldots, x_n$ be their respective INDEX values, and $e_0, \ldots, e_n$ be their respective EP values. Then in order for d_0 to count as a result of *fusing* $d_1, \ldots, d_n$, one of the following three conditions has to be satisfied:

a. $d_0 = d_1 = \cdots = d_n$.

b. $d_1, \ldots, d_n$ are alphabetic variants of each other,[7] $p_0 = p_1 = \cdots = p_n$, $x_0 = none$, and $e_0 = e_1 \oplus \cdots \oplus e_n$.

c. $p_0 = p_1 = \cdots = p_n$, $x_0 = x_1 + \cdots + x_n$, and for each i such that $1 \leq i \leq n$, the result of replacing all occurrences of x_i with $x_1 + \cdots + x_n$ in e_i is e_0.

Additionally, it is assumed that the option given in (41c) cannot be invoked when the structure involved is a disjunctive coordinate structure.

When multiple domain objects representing quantifiers are fused with each other in accordance with the condition in (41a) and right-node-raised, those domain objects, which are equated with each other, end up representing a single quantifier, and that quantifier therefore has to take wide scope over all the constituents out of which it has been right-node-raised, since otherwise some of the variables to be bound by the quantifier would remain unbound, in violation of the variable-binding condition. In the structure shown in Figure 13, for instance, the rightmost domain object in the order domain of the first daughter and the rightmost domain object in the order domain of the second daughter are fused in accordance with (41a), and the result is placed at the right edge of the order domain of the mother, creating a structure in which the universal quantifier expressed by the domain object obligatorily takes wide scope over the conjunction.

If the domain objects representing quantifiers are fused with each other in accordance with the condition in (41b) instead, then the structure as a whole is given essentially the same interpretation that it would have been given if RNR had not applied. In the case of example (36), this would lead to an interpretation that contains two instances of the universal quantifier.

In this account of example (36), the fact that the structure involved happens to be a coordinate structure plays no special role. As a consequence, the same account can be applied to an example like (37) as well, in which a quantifier has been right-node-raised out of a non-coordinate structure.

Let us next consider examples (38) and (39). In the theory proposed in Yatabe (2012), sentences like these result when domain objects that are right-node-raised out of different phrases are fused with each other in accordance with condition (41c). When condition (41c) is invoked, the newly created domain object is required to denote the sum of the denotations of the original domain objects. Thus, in the case of the sentence *John avoided and Bill ignored similar issues* in (38), in which the DP *similar issues* is right-node-raised out of two Ss, the domain object representing that DP in the order domain of the root node of the

---

[7] The term *alphabetic variant* is defined in Yatabe and Tam (2021).

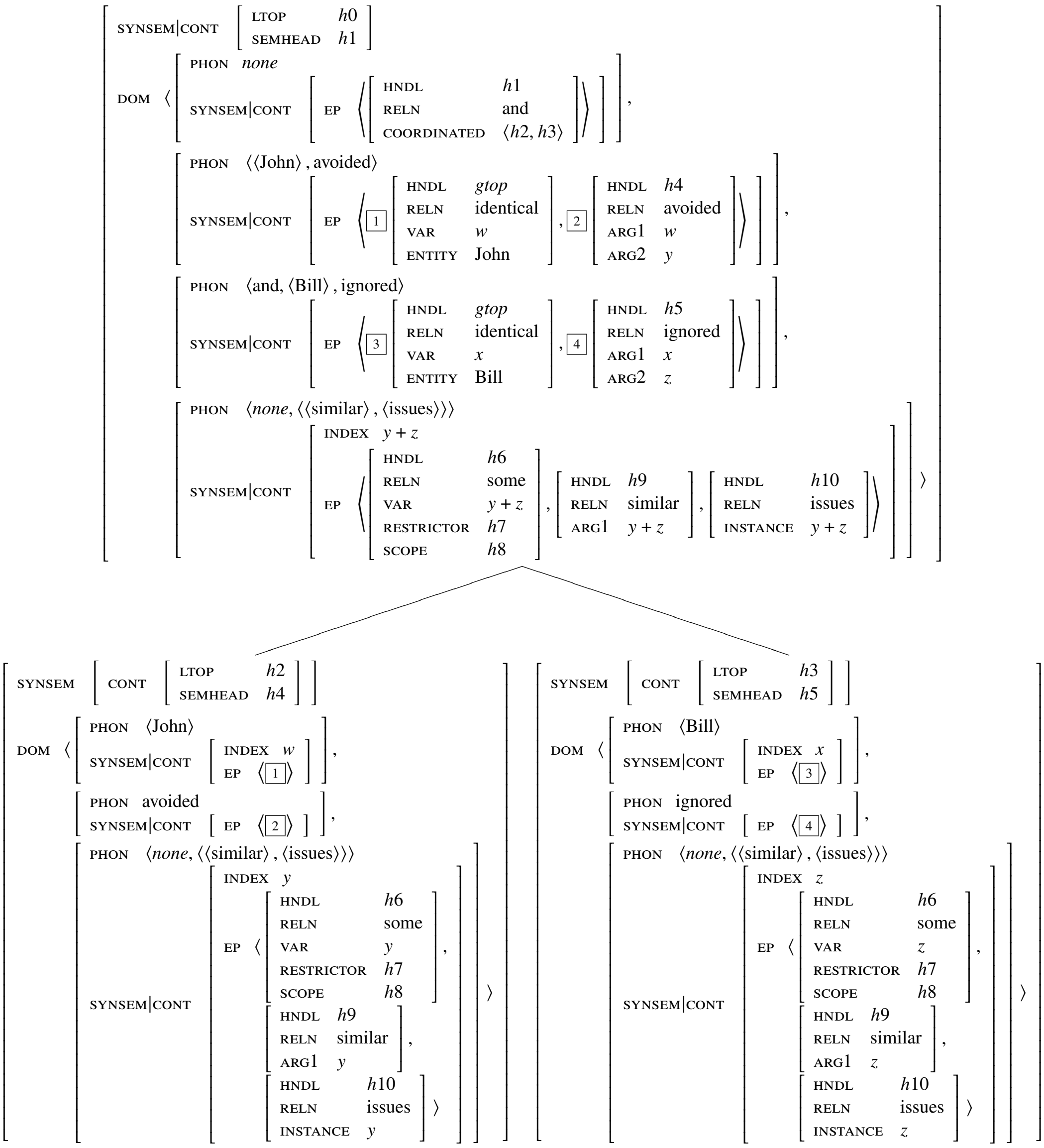

Figure 14: Part of a structure that can be assigned to example (38)

sentence, which is created by fusing the domain object representing that DP in the order domain of the first conjunct and the domain object representing that DP in the order domain of the second conjunct, denotes the sum of the entity that John avoided and the entity that Bill ignored, with the result that the entire sentence is correctly predicted to be able to mean "For some $x + y$ such that $x + y$ are similar issues, John avoided x and Bill ignored y".

Let us examine the way the sentence is assigned that interpretation in a little more detail. Figure 14 shows part of the structure assigned to the sentence on the assumption that there is an unpronounced determiner at the beginning of the DP *similar issues*, a determiner that contributes the meaning of an existential quantifier. What is shown in the figure is the local subtree in which the two Ss are conjoined and the grammatical object is right-node-raised. Recall that the meaning of a sentence is computed by first totally compacting the root node of the phrase-structure tree representation of the sentence. When the mother node of the local subtree depicted in Figure 14, which is the root node of the entire sentence, is totally compacted, we obtain the EP value shown in (42).

(42) $\langle\, h1 : \text{and}(h2, h3),\ \ gtop : \text{identical}(w, \text{John}),\ \ h4 : \text{avoided}(w, y),$
$gtop : \text{identical}(x, \text{Bill}),\ \ h5 : \text{ignored}(x, z),\ \ h6 : \text{some}(y + z, h7, h8),$

$h9 : \text{similar}(y + z),\ \ h10 : \text{issues}(y + z)\ \rangle$

The LTOP value of the NP *similar issues* must be equal to $h7$, the RESTRICTOR value of the *some* predicate, because of what the lexical entry of the unpronounced determiner says. The SEMHEAD value of the NP, which is required by (16d) to be identical to the LTOP value of *similar* and also to the LTOP value of *issues*, must either be locationally equivalent to or outscope $h9$ and $h10$, due to the LTOP-over-SEMHEAD requirement, which requires the LTOP values of *similar* and *issues* to either be locationally equivalent to or outscope their respective SEMHEAD values. Since the LTOP value of the NP must either be locationally equivalent to or outscope the SEMHEAD value of the NP, it follows that $h7 \geq h9$ and that $h7 \geq h10$. At the same time, the existential quantifier has to take scope over both conjuncts; otherwise there would be an unbound instance of either y or z in violation of the variable-binding condition, because the LTOP-over-SEMHEAD requirement dictates that the elementary predication "avoided(w, y)" must be part of the first argument of the *and* predicate and that the elementary predication "ignored(x, z)" must be part of the second argument of the *and* predicate. Therefore the meaning of the sentence is determined to be (43).

(43) $\text{identical}(w, \text{John}) \wedge \text{identical}(x, \text{Bill})$
$\quad\wedge\ \text{some}(y + z,$
$\qquad\qquad \text{similar}(y + z) \wedge \text{issues}(y + z),$
$\qquad\qquad \text{and}(\text{avoided}(w, y), \text{ignored}(x, z)))$

This representation expresses the intended truth conditions on the assumption that an elementary predication of the form "similar($x_1 + \cdots + x_n$)" is true if and only if the n entities denoted by $x_1, \ldots, x_n$ are similar to each other. (See Yatabe (2023) for more details on this assumption.)

This account for sentence (38) is applicable to example (39) as well, since here again the fact that the structure involved happens to be a coordinate structure plays no special role. The condition in (41c) can be invoked in any type of structure other than disjunctive coordinate structure.

Notice also that the theory has no problem dealing with an example like (44) below, where an expression has been right-node-raised out of a nested coordinate structure.

(44) Alex borrowed or Chris bought and Pat stole, a total of 20 books.

This HPSG-based theory compares favorably with other theories of RNR. First, theories that have been developed within the framework of Categorial Grammar (CG) do not generalize correctly to cases of RNR involving non-coordinate structure, which, as shown by the examples in (37) and (39), can affect semantic interpretation the same way that RNR out of coordinate structure can, pace Belk, Neeleman, and Philip (2023). The CG-based account of sentences like (37) proposed in Steedman (1996) serves as an illustration of the difficulty that non-coordinate RNR poses for Categorial Grammar. Steedman proposes to analyze sentences like (37) as the rightward-extraction counterparts of parasitic-gap constructions such as *articles which I will throw in the trash without reading*. Since in this analysis it is assumed that English has both leftward extraction and rightward extraction and that the two types of extraction constructions are generated by the same grammatical devices, it becomes necessary to add to the grammar some mechanism that accounts for the contrast between (45) and (46).

(45) Which island did you [travel to with Keats], and [return from with Chapman]? (from Steedman (1996))

(46) *Keats will [travel to with Chapman] the Isle of Capri. (from Steedman (1996))

Steedman thus goes on to propose an elaborate system of features that is designed to capture this contrast. As pointed out in Yatabe and Tam (2021, p. 40), however, the system in question turns out to be unable to deal with sentences in which a right-node-raised expression is simultaneously a complement of a preposition and a heavy-NP-shifted object of a transitive verb.

Second, theories of RNR developed within the framework of Transformational Grammar have been unable to provide a descriptively adequate account of the syntax and semantics of sentences like (38) and (39). In Transformational Grammar, RNR has been analyzed as resulting from deletion, movement, multidominance, or some combination thereof, but none of those analytical options is appropriate for sentences like (38) and (39). RNR in those sentences cannot be a type of phonological deletion because it is not meaning-preserving, cannot be movement because it does not induce rebracketing (see Belk, Neeleman, and Philip (2023) for a recent discussion on this point), and cannot be an effect of multidominance either because the right-node-raised expression needs to contribute a different variable to the semantic representation of each phrase out of which it is right-node-raised. To elaborate on this last point, an analysis in terms of multidominance entails, in the case of the example in (38), that the object of *avoided* and the object of *ignored* are the same node, and therefore is not compatible with the view, which seems at least reasonable, that the sentence needs to be associated with a semantic representation like (43), in which the second argument of the *avoided* predicate and the second argument of the *ignored* predicate are different variables.

And thirdly, unlike the other theories, the theory I am advocating extends naturally to a sentence like (47), which I assume is generated by right-node-raising out of each conjunct a relative clause that has been extraposed out of the subject DP.

(47) Every woman is smiling and every man is frowning who came in together.
(from Fox and Johnson (2016))

Assuming that a relative clause is given the same index as the nominal it modifies, the domain object representing the relative clause in the order domain of the first conjunct, which means something like "x came in together", and the domain object representing the relative clause in the order domain of the second conjunct, which means something like "y came in together", can be fused with each other in accordance with (41c) to produce a new domain object meaning something like "$x + y$ came in together", which can then be placed in the order domain of the root node of the sentence. Thus, given some additional assumptions about so-called resumptive quantification (see May (1989)), the sentence is correctly predicted to be able to mean "For every $x + y$ such that x is a woman, y is a man, and $x + y$ came in together, x is smiling and y is frowning" (see Yatabe and Tam (2021) for details of this account). An account like this is not available in theories in which semantic composition is performed through function application and relative clauses are assumed to denote functions. The account proposed in Fox and Johnson (2016) is in my view marred by its apparent reliance on the questionable assumption that a string like *woman man* can be a grammatical phrase synonymous with *woman and man*.

10 Concluding remarks

In this document, I have presented an HPSG-based theory of the syntax-semantics interface whose main components are (i) a DRT-like mechanism for semantic composition that relies on the assumption that predicates are interpreted as elementary predications rather than as functions, (ii) a system of semantic representations that allows underspecification of scope relations, and (iii) apparatus for compaction-driven meaning assembly, which implements the view that semantic composition is performed each time a larger *prosodic* constituent is created. Much of what I have done is recapitulation of material that has already appeared in the literature, but in section 6, I made a new theoretical proposal about how coordination of scopal predicates, including coordination of quantifiers, might be handled in MRS.

References

Bayer, Samuel Louis (1997). *Confessions of a Lapsed Neo-Davidsonian: Events and Arguments in Compositional Semantics*. Outstanding Dissertations in Linguistics. Garland Publishing Inc., New York & London.

Belk, Zoë, Ad Neeleman, and Joy Philip (2023). What divides, and what unites, right-node raising. *Linguistic Inquiry*, 54:685–728.

Copestake, Ann, Dan Flickinger, Carl Pollard, and Ivan A. Sag (2005). Minimal recursion semantics: An introduction. *Research on Language and Computation*, 3:281–332.

Fox, Danny and Kyle Johnson (2016). QR is restrictor sharing. In *Proceedings of the 33rd West Coast Conference on Formal Linguistics*, pages 1–16, Cascadilla Proceedings Project, Somerville, MA.

Gazdar, Gerald (1980). A cross-categorial semantics for coordination. *Linguistics and Philosophy*, 3:407–409.

Heim, Irene and Angelika Kratzer (1998). *Semantics in Generative Grammar*. Blackwell, Malden, Massachusetts.

Jackendoff, Ray (1977). *$\bar{X}$ Syntax: A Study of Phrase Structure*. MIT Press, Cambridge, Mass.

Kamp, Hans (1984). A theory of truth and semantic representation. In Jeroen Groenendijk, Theo M. V. Janssen, and Martin Stokhof, editors, *Truth, Interpretation and Information: Selected Papers from the Third Amsterdam Colloquium*. Foris, Dordrecht, pages 1–41.

Kathol, Andreas and Carl Pollard (1995). Extraposition via complex domain formation. In *33rd Annual Meeting of the Association for Computational Linguistics*, pages 174–180, Morgan Kaufmann, San Francisco.

Koenig, Jean-Pierre and Frank Richter (2021). Semantics. In Stefan Müller, Anne Abeillé, Robert D. Borsley, and Jean-Pierre Koenig, editors, *Head-Driven Phrase Structure Grammar: The Handbook*, number 9 in Empirically Oriented Theoretical Morphology and Syntax. Language Science Press, Berlin, pages 1001–1042.

Kubota, Yusuke and Robert D. Levine (2020). *Type-Logical Syntax*. MIT Press, Cambridge, Massachusetts.

May, Robert (1989). Interpreting logical form. *Linguistics and Philosophy*, 12:387–435.

Partee, Barbara H. and Mats Rooth (1983). Generalized conjunction and type ambiguity. In Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow, editors, *Meaning, Use, and Interpretation of Language*. Walter de Gruyter, Berlin, pages 361–383.

Pollard, Carl and Ivan A. Sag (1994). *Head-Driven Phrase Structure Grammar*. University of Chicago Press, Chicago.

Reape, Mike (1994). Domain union and word order variation in German. In John Nerbonne, Klaus Netter, and Carl Pollard, editors, *German in Head-Driven Phrase Structure Grammar*. CSLI, Stanford, pages 151–197.

Sag, Ivan A., Thomas Wasow, and Emily M. Bender (2003). *Syntactic Theory: A Formal Introduction*, second edition. CSLI, Stanford.

Steedman, Mark (1996). *Surface Structure and Interpretation*. MIT Press, Cambridge, Mass.

Steedman, Mark (2000). *The Syntactic Process*. MIT Press, Cambridge, Mass.

Williams, Edwin (1990). The ATB theory of parasitic gaps. *The Linguistic Review*, 6:265–279.

Yatabe, Shûichi (2001). The syntax and semantics of left-node raising in Japanese. In *Proceedings of the 7th International Conference on Head-Driven Phrase Structure Grammar*, pages 325–344, CSLI, Stanford.

Yatabe, Shûichi (2004). A comprehensive theory of coordination of unlikes. In *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar*, pages 335–355, CSLI, Stanford.

Yatabe, Shûichi (2009). Evidence for the linearization-based theory of long-distance scrambling in Japanese. In Linda Uyechi and Lian-Hee Wee, editors, *Reality Exploration and Discovery: Pattern Interaction in Language & Life*. CSLI, Stanford, pages 271–286.

Yatabe, Shûichi (2012). Comparison of the ellipsis-based theory of non-constituent coordination with its alternatives. In *Proceedings of the 19th International Conference on Head-Driven Phrase Structure Grammar*, pages 454–474, CSLI, Stanford.

Yatabe, Shuichi (2023). Further evidence for an HPSG-based theory of the semantics of *different* and *the same*. In *Proceedings of the 30th International Conference on Head-Driven Phrase Structure Grammar*, pages 143–163. https://doi.org/10.21248/hpsg.2023.8.

Yatabe, Shûichi and 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.