

Identification of Relative Clause as a Nominal Dependency Relation

Tanmoy Bhattacharya
Department of Linguistics
University of Delhi

Tanmoy@linguistics.du.ac.in

Nguyen Chi Duy Khuong
Department of Linguistics
University of Delhi

khuongvn2000@yahoo.com

ABSTRACT

This paper proposes to look at the Relative Clause construction in Vietnamese with respect to its status as an adjunct (the dominant view of Chomsky (1971)) as well as a complement, a recent view put forth in Kayne (1994). In addition, we propose to provide a simpler solution for RC recognition which overcomes the need to identify first an RC structure as either a complement or an adjunct. In effect, the paper bypasses the need to feature specification on either the adjunct or the head by showing that there is in fact no adjunct generated in the construction of a RC.

1 Introduction

In the realm of constructions involving adjuncts, formalisms adaptable to implementation (for example, HPSG) have been beset with the problem of first identifying the adjunct and then determining the place in the tree where that adjunct belongs. In short, finding an algorithm to determine a suitable *addressing technique* whereby an adjunct structure is integrated in the main body of the clause, has eluded researchers in this field. This paper highlights this problem further from the perspective of Relative Clause (RC) construction in Vietnamese and investigates the possibility of gaining insight, if any, from current theories of RC construction within the Principles and Parameters theory with the view of getting around the problem of the status of adjuncts.

2 Classic HPSG: The duality of representation

The problem outlined above faced while generating/ parsing an RC structure, leads invariably to a type of solution that unsatisfactorily proliferates features required in

producing a syntactic parse. The issue highlighted here is not a problem of these formalisms alone. In semantics work in general, and in Categorical Grammar in particular, it is assumed that the complement/ adjunct reduces to whether the adjunct is a functor or an argument. The head-complement idea that this assumption incorporates is conducive to the spirit of HPSG type formalism in general where it is easier to see complements as semantic arguments of their heads.

Adjunct selection is different from complement selection in at least two ways:

- (a) The range of categories that an adjunct can modify is much broader than the range modified by the complements, and
- (b) The number of adjuncts that a category can be modified by is not pre-fixed.

Given this, Pollard and Sag (1987) [PS] considers the option that the specification responsible for identifying an adjunct-main clause dependency is not be specified on either the adjunct or the head but be determined by rules of grammar. With respect to RCs, for example, it is assumed that there is a rule that in some way combines Ns and RCs. The solution, however, rests on specifying (assuming) the hierarchy of types/ subtypes (head-structure and head-adjunct-structure respectively).

The classic HPSG solution to the problem of identifying the adjunct, especially in relation to an RC, is thus sorted out by specifying a feature like [RELCLAUSE] initially, as shown in Figure 1 in Appendix.

In this sign, there is no obvious scope of unification of the ADJ-DTR structure with the HEAD-DTR structure part from begging the same lakh rupees question: who selects whom? Apart from this, the full range of head-adjunct pairs will require many such rules. The formal implementation of the RC selecting the head, though attractive, is very hard to achieve (as it is in a non-constraint based framework); the alternative, of allowing the head to select for their adjuncts is more implementable. The main rationale for opting for this direction (from Head

to Adjunct) is that the range of syntactic projections that are modified by a certain adjunct is predictable and highly limited. Thus, RCs modify the projection of category NFORM.

However, the way this is adopted within the classical HPSG model is to assume that each lexical sign specifies a value for the head feature ADJUNCTS. Thus every common noun is assumed to bear the specification [SYN| LOC| HEAD| ADJUNCTS {RELCLAUSE, ...}] and subjected to the head-adjunct rule shown in Figure 2 of the Appendix.

3 Revised HPSG

The duality of representation problem raised in this paper so far is addressed to some extent in Sag (1995) who presents a slightly modified version of the classic model which provides a cumbersome procedure of characterising NONLOCAL feature inheritance in terms of the features INHERIT and TO-BIND in addition to the NONLOCAL feature principles proposed. In Sag (1995) it is claimed that only the highest verb of the RC is morphologically distinguished. The way to account for this in HPSG is straightforward as the highest verb (rather than a null relativizer) heads the clause. Since all modifiers bear the feature MOD of the HEAD, a RC bearing [MOD N] will percolate up to the highest verb. This is shown in (22) of the Appendix for Korean which employs the participle strategy of RC formation.

However, this claim that the highest verb in an RC shows up RC-related morphology will not simply work for languages that show no RC morphology, more common in isolating languages like Vietnamese but also, in Japanese (another relative participle language) and in well-studied languages like English which do not show any special morphology on the verb inside the RC.

Consider the following Japanese RC:

(1) Watashi ga inu ga
 I NOM dog NOM
 taberu ring-o miru
 eats apple-ACC see
 'I see the apple which
 the dog eats.'

The traditional structure for this RC is shown in (23) of the Appendix. In this example the highest verb *miru* 'see' does not bear any relative morphology (the well know reason for reanalysis

in the parsing tree for such sentences, see Sharma and Bhattacharya, 2002, Sharma, 2004).

Similarly, in the following Vietnamese example, the embedded verb *tang* 'offer' does not show any special relative morphology:

(2) kwyen sak [ma ban tang
 CLA book that you offer
 cho toi] rat hay
 for me very good
 'The book that you
 offered me is very good.'

Sag (1995) addresses this problem by first of all considering that words are subject to a constraint that defines their SLASH value in terms of the SLASH values of their ARGUMENT-STRUCTURE list, this is shown in Figure 3 of the Appendix.

However, this is not enough to recognize the RC and a HEAD feature REL is required which takes a set of referential indices as values. Thus the relative pronoun *who* in English is specified as in Figure 4 of the Appendix.

The REL feature is passed up from the *Wh*-word via the heads that select it to the phrase that directly dominates it. At the top level of an RC, then, the grammar can simply impose the requirement that the non-head daughter has a non-empty REL specification which will be sufficient to guarantee that the non-head daughter has a *Wh*-word somewhere within it. A *Wh* Inheritance Principle ensures the passing up of the REL feature and a *Wh*-RC furthermore is subject to another constraint which ensures that the non-head daughter of the RC must have the same index as the MOD value of the RC.

3.1 Types/ sub-types

However, the issue of selection of a feature of a construction (REL here) is assumed to be derived from a set of hierarchies that both the classic and revised version of HPSG take to be primitive. So, for example, clauses are distinguished from non-clauses and clauses are further sub-divided into at least four sub-types: *decl(arative)-cl(ause)*, *inter(rogrative)-cl(ause)*, *imp(erative)cl(ause)* and *rel(ative)-cl(ause)*. Apart from this, the heads are also divided in terms of their position in the tree and their relation with the rest of the structure. So a particular phrase is cross-classified in terms of its clause type as well as the head type it is associated with.

Furthermore a *rel-cl* subtype is associated with the constraint shown in Figure 5 of the Appendix. The constraint [MC ___] ensures that

RCs are not main clauses and the [MOD [HEAD *noun*]] ensures that any RC introduced in the *head-adjunct-phrase* will modify a nominal head daughter.

In order to formulate a rule that combines nouns and RCs (see section 2), it is assumed that the constituent-structure *type headed-structure* has a *subtype head-adjunct-structure*. Thus, in the classical version, referring to Figure 1, the fact that we know that the non-head daughter is an adjunct daughter (ADJ-DTR) is by specifying the *type/ sub-type* of the non-head daughter.

3.2 Semantics of the Gap

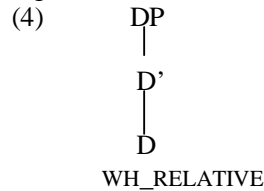
This section shows that in dealing with the semantic input structure for generating/ parsing a RC structure, the problem of unsatisfactory feature proliferation surfaces in the same way as in producing a syntactic parse. What about the semantics of the RC? Note, first though, that HPSG claims to integrate the semantic information of a structure as well. However, there is a finer distinction, the gap left by the operator in a RC construction is famously known to be semantically problematic in the sense that it represents an individual but there is no suitable individual to bind it. In fact, this problem is considered to be one of the motivations for introducing the notion of a variable in semantics and, in turn, in syntax (more on this Heim and Kratzer 1998). It needs to be pointed out that such a finer semantic distinction is not captured in the semantics that a feature like [CONT] can represent. So we come back to the question of representing the semantics of the RC.

A ready and rough approach to the semantic import of the RC can be seen in a simple system such as that of Hunter (2004) who presents a general generation procedure that requires two semantic input structures, one for the main clause and the other for the RC, for parsing of a RC construction. Consider the following in this connection:

(3) I eat the apples which
the men bought.

First, one input is needed for the relative clause, *which the men bought*, this is shown in Figure 6 of the Appendix. Here in this template a new argument is introduced, *relativewh*. This argument is inserted in the position of the nominal which the RC modifies; thus the RC modifies *apples*, which has the theme role in the buying event.

The DPs generated for *relativewh* arguments have a new type of determiner:



This DP is inserted in the position dictated by the theta grid of the English verb *buy*, and the presence of a WH_RELATIVE argument causes a WH feature to be encoded on the C.

Looking at the main clause of (3), in order to link the two clauses, an identifier *subclause0* referring to the structure in it is further introduced. The input to represent the main clause is as shown in Figure 7 of the Appendix.

Yet another type of argument type has been introduced here in Figure 7, *relative*. This is used for arguments in a main clause which are modified by a relative clause. The final part of the information in this new argument links the argument to the semantic input for the relative clause which modifies it. The DPs constructed for relative arguments are similar to *literal* arguments, but the CP generated from the semantics of the relative clause is *added as an adjunct* to the NP.

In this system, the important problem that challenges any system attempting to work out the parse for a RC, that is of identifying the status of adjunct, is bypassed by simply by assuming that the CP generated “is added as an adjunct”. The other aspect of the problem raised in this paper with regards to these formalisms is the recurrent theme of the duality of representation, one for the main clause and one for the RC.

4 Status of adjunct within P&P theory

We wish to point out that in the current debate surrounding the status of adjuncts it is precisely the dualism pointed out above with respect to the functor/ argument nature of adjuncts that is brought forth in theories of Cinque (1999) and Ernst (2000) (see Thangjam, 2004, in preparation) for more on this debate)¹.

The various questions that arise in relation to the status of adjuncts are directly relevant to the concerns of the present paper. These are as follows:

- (i) How is adjunct selection characterised?

¹ However, PS disfavours the functor role of adjuncts as it appears to be in contrast to the HPSG formalism.

- (ii) Do heads select for adjuncts or vice-versa?
- (iii) What principles determine the order of adjuncts?
- (iv) Can they be sisters to complements?
- (v) How should adjuncts be introduced by rules of grammar?

4.1 Relative Clauses within P&P

In the realm of RCs, the debate translates into the adjunction versus selection of the RC. The former, more popular view of RC is that which derives from Chomsky (1977) where the RC is considered to be adjoined to the relativized NP. Thus in (5) the RC *that John made* is adjoined to the base-generated head noun *claim* which is linked with a *Wh*-Operator in [Spec,CP] (or a real *Wh*, if there is one as in *Wh*-relatives) by an interpretive relation (binding, predicative or 'construal'):

(5) [_{DP} the [_{NP} [_{NP} claim_j]
 [_{CP} OP_j that John made t_j]]]

However a restrictive RC is interpreted within the scope of a Determiner, especially when the Det is a Quantifier, as in the following:

(6) Every girl that Mary saw
 $\forall x$ [girl(x) ^ Mary saw(x)]

Furthermore there is good evidence to show that at LF the relative head undergoes reconstruction to its pre-movement position; only after reconstruction the right C-command configuration achieved:

(7) ~~The interest in each~~
~~other_i~~ that John and Mary_i
 showed interest in each
 other_i was fleeing.

(Jackendoff, 1972,
 Schachter, 1973)

The adjunction analysis (as sketched in (5)) cannot account for the C-command requirement:

(8) [_{DP} The [_{NP} picture of
 himself]_i [_{CP} OP_i that John_i
 painted t_i]]

Thus the *Head Raising* account was proposed as an alternative to the adjunction analysis of RC. In this account, the head undergoes raising from the relative clause internal position to a higher position in the domain of the CP (most typically to [Spec,CP]). This position has been argued for quite extensively in the generative literature (Brame 1968, Schachter 1973, Vergnaud 1974, Kayne 1994, among many others). A typical view of the RC in the Head

Raising account is as shown in (24) of the Appendix.

4.2 Evidence for the complement structure of RCs in Vietnamese

In this section, we show that there are at least two kinds of empirical evidence in Vietnamese in support of the Kaynean structure of RC (as shown in (24)).

First, it is observed that there is a selection relation between a Det and the related RC in terms of definiteness/ indefiniteness nature of the Det:

- (9)a. She is that kind of person.
- b. She is **the** kind of person *(that is always complaining).
- (10)a. He did it in that way.
- b. He did it **the** way *(that annoyed me).

A similar type of restriction is observed in the following data from Vietnamese:

- (11)a. *nguoì dan-ong*
 CLA man
 [_{RC} *dang ngoì o dang-kia*]
 ING sit in there
la anh trai kua toi
 is brother boy of me
 'The man sitting over there is my brother.'
- b. **nguoì dan-ong dang ngoì o dang-kia*

Thus, (11b) shows that omitting the RC is not an option. Similarly for the following:

- (12)a. *ba-ta la mot ngoùì*
 She is a CLA
phu-thuy
 evil
 'She is an evil.'
- b. *ba-ta la mot loai*
 she is a kind
phu-thuy *(*chunyen*
 evil eat
 an thit heo)
 only meat pig
 'She is the kind of evil person who only eats pig's meat.'
- (13) *Anh-ta thuong hoi*
 he often ask
theo
 according

kai loi *(ma lam ban
 CLA way that make you
 buk-minh)
 annoy
 'He often asks in a way
 that makes you annoyed.'

In both (12) and (13), the RC is obligatory, showing that the definiteness of the noun associated with the RC is somehow responsible for this obligatoriness. In other words, the Classifier element (CLA) which endows the noun with definiteness has a selectional relation with the associated RC, in other words, Kayne's structure for RC makes the right predictions for Vietnamese RCs.

The data from Reflexive Binding show similar reconstruction effect as in (7) above:

(14) buk *tranh ve chinh*
 CLA picture about SELF
 anh-ta ma John ve, ...
 him/ he that J drew
 'The picture of himself
 that John drew.'

The reconstruction of the RC *tranh ve chinh* to its base position alone can explain the C-commanding configuration between the antecedent and the reflexive.

4.3 Head Raising Analysis of RC in Vietnamese

Having shown in the previous section the obvious selectional connection between the CLA and the RC in Vietnamese, we proceed now to look more carefully at the structure of RCs in this language. Consider the following example in this connection:

(15) *nguai dan-ong ma ban*
 CLA group man that you
 gap hom-qua
 meet yesterday, ...
 'The man that you met
 yesterday...'

For this particular RC, there is nothing to choose between either the adjunction or the complement analysis of RC as there is no clinching evidence in favour of either. However, consider now the following where it is possible to optionally drop the relative COMP *ma* 'that':

(16) *kau be (nguai) (ma)*
 boy small CLA that
 thi rot tuna qua la
 examine fail week
 ban qua nam
 last is friend of Nam

'The small boy who
 failed last week is
 Nam's friend.'

A similar kind of asymmetry in retaining or dropping of *ma* is observed in appositives as well:

- (17)a. *Nam, nguai ma ban gap*
 Nam, CLA that you meet
 hom-qua
 yesterday
 'Nam, whom you met
 yesterday.'
 b. *Nam, nguai (#ma)dang*
 Nam CLA #that ING
 dung o dang-kia ...
 stand in there
 'Nam, who is standing
 over there,...'

The asymmetry observed above can be summarised as follows. For object relatives, *ma* 'that' is obligatory whereas for subject relatives, it is either optional or downright bad. We will show that this state of affairs can only be accounted for in a complement structure of the RC and not in the adjunction analysis.

The analysis proffered here is a novel one in the sense that although the Kaynean structure of the RC is seen to be the most appropriate for Vietnamese RCs, the crucial issue that clinches the argument is based primarily on a certain economy principle of the Minimalist Program (MP) (Chomsky 1995).

4.4 Minimalist Economy Principles and the Kaynean Structure of RCs

In this section we show that the following Economy Principle of the MP provides crucial support for a Head Raising analysis of the RC in Vietnamese:

- (18) Equidistance
 If α , β are in the same
 minimal domain, they
 are equidistant from γ

In other words, two targets of movement are equidistant if they are in the same minimal domain. This is illustrated briefly in (25) of the Appendix.

The chain, CH_1 , created by the movement of Y to X defines the minimal domain as follows:

- (19) $MIN(CH_1) = \{Spec1, Spec2, ZP\}$

The definition of Equidistant as in (18) implies that now (i.e., after the movement $Y \rightarrow X$) Spec1 and Spec2 are *equidistant* from ZP and anything that ZP dominates. The result of this is that now

the raising of ZP can in effect skip the Spec2 position, although Spec2 is the first legal landing site.

With this in mind, let us look at the following pair of examples:

- (20)a. *kanh hoa ma John mua*
 CLA flower that bought
 'The flowers which
 John bought.'
 b. *kanh hoa rot tu*
 CLA flower fell from
kai kay nay
 CLA tree this
 'The flower which fell
 down from the tree.'

A head raising analysis structure for (20a) is as in (26) of the Appendix.²

The Head Raising Analysis demands the movement of the head noun (in this case *hoa* 'flower') to the [Spec,CP] position. However, note that the movement of the DP *hoa* across the DP *John* would violate the Minimal Link Condition (MLC), defined as follows:

- (21) MLC
 α can raise to K only
 if there is no
 legitimate operation
 Move β targeting K,
 where β is closer to K.

Although Kayne (1994) is not concerned with MLC since the only motivation for the Head Raising analysis for him is empirical (that the D and its complement do not form a constituent), we claim that Kayne's analysis tightened by Minimalist economy conditions in general,³ can only be for the better.

In connection to (26), notice that for Equidistance to apply in order to allow the movement of the DP to [Spec,CP] shown, it must be the case that some other head has moved to C prior to the raising of the relative head. We assume that the fact that the R-COMP *ma* is realised only in the case of object relatives has to do with this other head (or a feature of the head) moving to C to license the R-COMP. In the case of subject relatives, the economy principle of

Equidistant need not be evoked as movement of the subject DP to [Spec,CP] does not have to skip a Spec position in between, thus predicting that the R-COMP head need not be realised in the case of subject relatives; the prediction, as we have seen before, is borne out.

5 What's in it for Parsing?

As we saw earlier in sections 1 and 2, that a computational implementation of structures involving adjuncts is beset with the crucial problem of a lack of a proper addressing technique, i.e., the formalism is undecided first of all about the issue who selects who (the head the adjunct or vice-versa?) and second with regards to the duality of featural representation at both ends (adjunct and main clause along with a plethora of constraints) in order to achieve a unification.

At this stage of the elaboration of the structure of the RC in Vietnamese, we would like to draw attention to the complete *absence* of the notion of adjunct in the formalism that we have adopted following mainly Kayne (1994). In other words, in terms of the formalism of HPSG, we have completely gotten rid of the need to include the ADJ feature in the specification of any rule to parse a RC. Consequently the issue of determining the appropriate place for incorporating an RC does not arise in this model. This we would claim is a definite advantage over formalisms that are bound to overload any algorithm by including features to identify and then place an RC in the main clause.

² The structure deliberately fudges various details to do with the status of the Classifier in such languages, but see Bhattacharya (1999) and Khuong (2004) for a detailed elaboration of the structure of the DP of languages with classifiers.

³ See Bhattacharya (2002, 2005) for a such a marriage between Kayne's Antisymmetry Theory and the MP.

References

- Bhattacharya, Tanmoy. 1999. *The Structure of the Bangla DP*. Ph.D., University College London.
- Bhattacharya, Tanmoy. 2002. Peripheral and Clause-internal complementizers in Bangla: A Case for Remnant Movement. *Proceedings of Western Conference in Linguistics 2000*, 100-112, Fresno, CA.
- Bhattacharya, Tanmoy. 2005. The Myth of Areal Linguistics: A View from the Interface, talk delivered at Nanzan University.
- Brame, Michael. 1968. A new analysis of the relative clause: evidence for an interpretive theory. Ms. MIT
- Chomsky, Noam. 1977. On Wh-Movement. *Formal Syntax*, edited by Peter Culivoer, Thomas Wasow and Adrian Akmajian, 71-132, New York: Academic Press.
- Chomsky, Noam. 1995. *The Minimalist Programm*. Massachusetts: CA, MIT Press.
- Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistics perspective*. Oxford University Press, Oxford.
- Ernst, Thomas. 2000. *The Syntax of Adjuncts*. The Cambridge University Press, Cambridge, UK.
- Heim, Irene and Angelika Kratzer. 1998. *Semantics in Generative grammar*. Blackwell Publishers, Massachusetts and Oxford.
- Hunter, Tim. 2004. *A Scalable Natural Language Generator*. Undergraduate thesis (B.E.), University of New South Wales.
- Kayne, Richard. 1994. *The Antisymmetry of Syntax*. MIT Press.
- Khuong, Nguyen C.D. 2004. *The DP in Vietnamese*. MPhil thesis, University of Delhi.
- Pollard, Carl and Ivan sag. 1987. *Information-Based Syntax and Semantics*. Stanford: CSLI.
- Sag, Ivan. 1995. English relative Clause Constructions. *Journal of Linguistics*.
- Sharma, Atreyee. 2004. *Construal and Reanalysis in Parsing Assamese Sentences*. MPhil thesis, University of Delhi.
- Sharma, Atreyee and Tanmoy Bhattacharya. 2002. The processing of empty categories in relative clause attachment in Assamese. Paper presented at the 24th LSI, AICL, Mysore.
- Schachter, Paul. 1973. Focus and relativization. *Language* 53:19-49.
- Thangjam, Hindustani Devi. 2004. Interfaces and Clause Structure. Paper presented at the 26th LSI, AICL, NEHU, Shillong.
- Thangjam, Hindustani Devi. (In preparation) *Interface Strategies and Adjuncts*. Ph.D. dissertation, University of Delhi.
- Vergnaud, Jean-Roger. 1974. French relative clauses. Ph.D. dissertation, MIT, Cambridge, Mass.

Appendix: Figures and Tree Diagrams

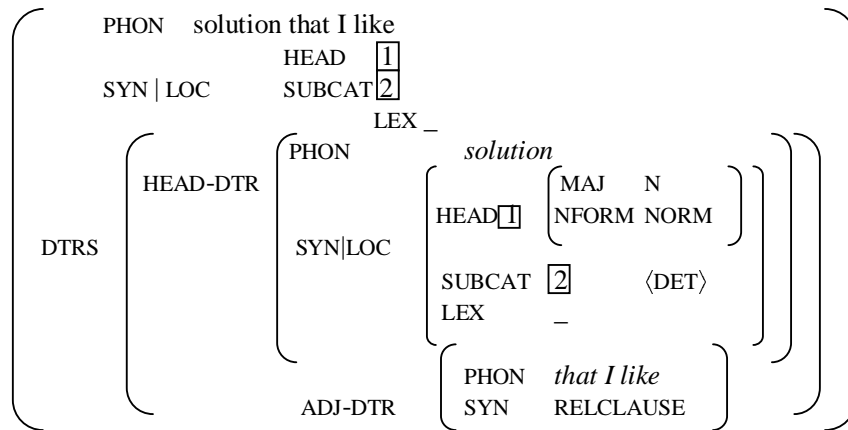


Figure 1: The Classic HPSG Sign for the RC *solution I like*

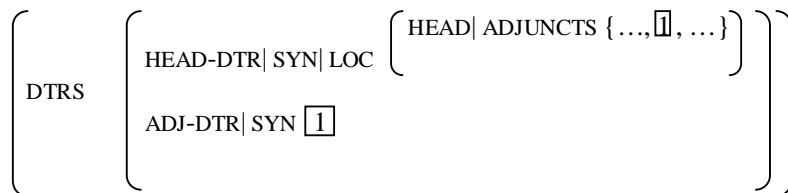
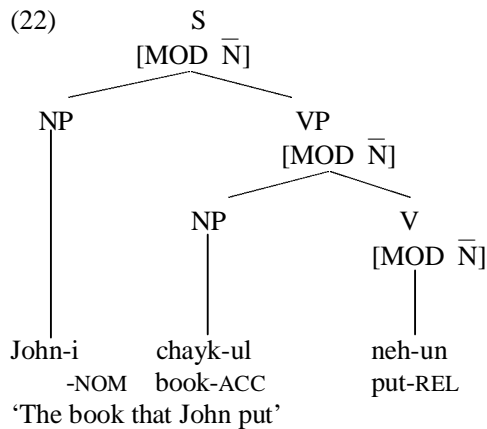
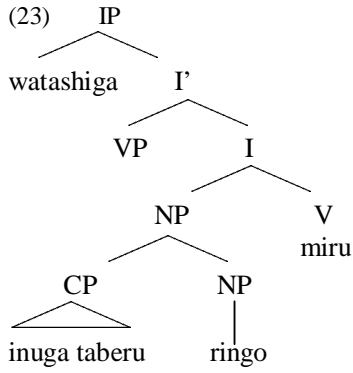


Figure 2: Classic HPSG representation of a Head-Adjunct rule





$$word \Rightarrow \left(\begin{array}{l} \text{BIND} \quad \boxed{0} \\ \text{ARG-ST} \quad \langle [\text{SLASH} \boxed{1}], \dots, [\text{SLASH} \boxed{n}] \rangle \\ \text{SLASH} \quad (\boxed{1}^{\cup} \dots^{\cup} \boxed{2}) _ \boxed{0} \end{array} \right)$$

Figure 3: The constraints that words are subject to in Sag (1995)

who (relative)

$$\left(\begin{array}{l} \text{CAT} \quad \text{NP} \\ \text{CONT} \quad [\text{INDEX} \boxed{3}] \\ \text{REL} \quad \{\boxed{3}\} \\ \text{QUE} \quad \{ \} \end{array} \right)$$

Figure 4: Specification of the Relative Pronoun in Sag (1995)

$$\left(\begin{array}{l} \text{HEAD} \quad \left(\begin{array}{l} \text{MC} \quad \text{---} \\ \text{INV} \quad \text{---} \\ \text{MOD} \quad \left(\begin{array}{l} \text{---} \\ \text{HEAD } \textit{noun} \end{array} \right) \end{array} \right) \\ \text{CONTENT} \quad \textit{proposition} \end{array} \right)$$

Figure 5: The constraint identifying a *rel-cl* sub-type

BUY	
AGENT	literal (MAN, PLURAL, DEFINITE)
THEME	relativewh
tense	PAST
negative	FALSE
question	FALSE
voice	ACTIVE

Figure 6: Input for the RC *which the men bought*

EAT	
AGENT	conversant (SINGULAR, +SPEAKER, -LISTENER, HUMAN_MASC)
PATIENT	relative (APPLE, PLURAL, DEFINITE, subclause0)
tense	PRESENT
negative	FALSE
question	FALSE
voice	ACTIVE

Figure 7: Input to the main clause using identifier *subclause0*

