Conflict and Reconciliation in Feature Structures*

Incheol Choi
(Kyungpook National University)

Choi, Incheol. 2011. Conflict and Reconciliation in Feature Structures. Linguistic Research 28(2), 311-328. The HPSG framework authorized by Pollard and Sag (1994) requires that a feature structure be totally well-typed and type-resolved (Carpenter 1992). In this paper, I showed that such tight feature logic poses various problems when it is implemented to deal with natural language phenomena. Building on the underspecification approach in Sag (2003), I proposed an HPSG approach which lets us avoid such problems not only in coordination but also in various extraction constructions. While easing the rigidity of the feature regime originally adopted in HPSG, this approach still maintains the rigorous mechanism of feature logic in that all possibilities of type resolution and feature membership are constrained in the type hierarchy and the lexical entries. (Kyungpook National University)

Key Words feature structures, type-resolution, well-typed feature structures, case, coordination, parasitic gaps, underspecification

1. Introduction

Pollard and Sag (1994) declares that a linguistic theory should be rendered in a formal logic as long as its content can be made clear and unambiguous in natural language. This line of doctrine can actually be driven from the one advocated by Chomsky (1957), in which it is said that only the formalized theory can provide solutions for many linguistic problems and recognize the productive potential. For the precise description of linguistic entities and eventual modeling of the natural language, the HPSG framework employs a system of typed feature structures.

A feature structure is a way of representing linguistic information of a linguistic entity and consists of features and their values. In HPSG framework, a linguistic

* I am grateful to Jong-Bok Kim for the detailed comments and discussion. I was also greatly benefited by anonymous reviewers. This research was supported by Kyungpook National University Research Fund, 2008.
entity is of a certain type. It means that the entity allows only a certain feature structure that is appropriate for it. For example, only the feature structure in (1a) between those in (1) is well-typed.

\[(1) \text{ a. } \begin{cases} \text{noun} \\ \text{CASE case} \end{cases} \quad \text{b. } \begin{cases} \text{noun} \\ \text{AUX boolean} \end{cases} \]

The feature structure in (1a) is equipped with the information that a nouns requires, for example, the case specification such as nom or acc. On the other hand, the AUX feature is not appropriate for nouns and hence the feature structure in (1b) is not well-typed.

The HPSG framework authorized by Pollard and Sag (1994) adopted further criteria of completeness. That is, they suggest that a feature structure be totally well-typed and type-resolved (Carpenter 1992). In a totally well-typed feature structure, every feature that is appropriate for the type should be present.

\[(2) \text{ a. } \begin{cases} \text{synsem} \\ \text{CATEGORY category} \\ \text{CONTENT content} \\ \text{CONTEXT context} \end{cases} \quad \text{b. } \begin{cases} \text{case} \\ \text{CASE nom} \end{cases} \]

For example, the linguistic entity \textit{synsem} can specify category, content and context information to be well-typed. If it omits any of these features, it will not be a well-typed feature structure. In addition, the value of a terminal feature should be maximally specified. A feature may have a value that is further specified by another feature structure or that is an atomic type. For example, the feature CASE requires the value \textit{case} as an atomic value without any further feature structure specification. In this case, the value should be a maximal type such as \textit{nom} or \textit{acc} but not the super type \textit{case}. This is because a feature structure should be type-resolved, following the tenet of Pollard and Sag (1994) (also King 1989, 1994).

The aim of this paper is to examine whether this type of strong version of feature logic can survive intact when it deals with natural language. Recent

---

1 \textit{synsem} is a linguistic entity that specifies the syntactic, semantic, and contextual information of a sign.
developments in HPSG framework, particularly Sag (2003), tend to admit that a structural description may not be in one-to-one relation with the feature structure. Following Sag’s retreat from the type resolution of the feature structure, I will show that such relief is inevitable to deal with various phenomena particularly regarding coordination, case mismatch and mixed category phenomena in English and Korean. In addition, I will provide a demonstration of the solution for the phenomena while adopting the Sag’s underspecification analysis.

2. Problems of the type–resolved feature structure

Traditionally, problems of the tight feature logic take place when it comes to coordination constructions.

(3) a. Pat is [wealthy and a Republican]. (Sag et al. 1985)
   b. Kim [likes bagels and is happy]. (Sag 2003)

As pointed out by Sag (2003), the bracketed coordinate structures in (3) do not maintain the parallelism of coordination. According to the strong version of the coordination principle in Pollard and Sag (1994), the CATEGORY and NONLOCAL value of each conjunct daughter is identical to that of the mother. However, the conjuncts in the bracket in (3a) each have a different part of speech and those in (3b) have different AUX values. To avoid this dilemma, Pollard and Sag (1994) propose a weak version of coordination principle, in which the CATEGORY and NONLOCAL value of each conjunct is subsumed by that of the mother. In this approach, the coordinate mother phrase in (3a) would only specify PRD + feature without specifying specific category such as adjective or noun. In addition, the whole coordinate phrase in (3b) does not specify the AUX feature. These types of underspecification not only violate the condition of type resolution, but also neglect the requirement that feature structures be type-resolved.

The tight feature logic also causes problems for parasitic relatives (Levine, Hukari and Calcagno 2001).

(4) Robin is someone who even good friends of t believe t should be closely watched.
In (4) the parasitic gap inside the relative clause shares the local value with the subject gap. However, then, the case values of the two gaps are in conflict and do not match with the relative pronoun who. To solve this problem, Levine, Hukari and Calcagno (2001) suggest that the relative pronoun who has a category that is cross-classified by nominative and accusative, hence satisfying any selectional requirement for the two cases. This account can work only when the selectional requirement imposed to the subject gap and parasitic gap involves non-maximal CASE values. For example, the case given to the subject gap should be compatible not only with the nominative but also with the mixed case nom-acc. Accordingly, this approach violates the requirement that a feature structure should be type-resolved.

A similar problem is raised with respect to Korean nominalization constructions.

(5) na-nun cheolswu-ka [pap-to mek-ess-ko
                I-TOP Cheolswu-NOM meal-even eat-PST-CONJ
                hakkyo-e-do kass-um]-ul an-ta
                school-LOC-even went-NML-ACC know-DECL
            “I knew that Cheolswu ate meal and went to school.”

The whole coordinate structure in the bracket (5) should be a noun phrase in that it is marked with the accusative case –ul. However, the first conjunct is headed by a verbal stem mek-ess- ‘ate’ which in turn combines with the verbal coordination affix –ko. The dilemma here takes place because the category of the first conjunct is not identical to that of the second conjunct. Choi (1999) tried to solve this problem by granting head status to the second conjunct and categorizing the um- attached word as a mixed category VN (see also Kim 2001).

(6) Verbal Nominal

    verb VN noun

In this approach, VN in the type hierarchy (6) is cross-categorized by two super types Verbal and Nominal and inherits all the constraints and feature structures from
Nominal and Verbal. Specifically, the um-attached word in (5) combines with the accusative marker -ul since it is VN, which is a subtype of Nominal. The second conjunct also takes a verbal argument since it is a subtype of Verbal. Above all, the categories of the first conjunct and the second conjunct are not identical but are in subsuming relation. This is because the first conjunct is categorized as the super type Verbal without further specification of the category. Therefore, in Choi’s approach, the category value of the whole coordinate phrase becomes VN, owing to the unification process between the underspecified super type Verbal and the maximal type VN. Although Choi (1999) did not mention the problem, it surely neglected the requirement that feature structures should be maximally resolved.

3. Previous Approaches in HPSG framework

The tight feature logic adopted in Pollard and Sag (1994) does not always seem to be harmonious with the natural language phenomena. In HPSG framework, the relaxation of the doctrine mainly comes from the requirement for the maximal resolution of the value. Such relaxation can be categorized by two types. The first type is to allow the underspecification of the value when an entity is inserted into a syntactic domain. In this type of approach, once the entity is inserted in the syntactic domain, its maximal value is decided by the phrasal schemata or constraints.

A good example of such analyses comes from Kim and Choi’s approach to Korean case assignment (2004).
The lexeme chayk-to ‘book-even’ specifies gcase for the value of the feature CASE. Since gcase is a supertype of nom and acc, it can undergo unification with either acc or nom. As illustrated in the structure in (7), the maximal case value acc is syntactically assigned following the HEAD-COMPLEMENT rule suggested in Kim and Choi (2004). The benefit of this type of approach is to restrict the underspecification of value to the lexical level or word.

On the other hand, Sag (2003) suggests a rather unconventional approach in that it allows the underspecification even in the syntactic domain. In his approach, feature structures should still be totally well-typed. However, the values of the features need not be maximal types in both lexical and syntactic domains. For example, in his approach, the coordinate phrase wealthy and a Republican in (3) does not specify a leaf of the type hierarchy in (8) but a nonmaximal type nominal regarding CASE feature.
(8) The type hierarchy for part of speech in Sag (2003)

```
   pos
  /   \
nominal verbal
  |
noun adj  prep verb
```

Specifically, Sag (2003) uses a formula for bounding constraints to permit multiple resolution.

(9) a. wealthy: $[\text{HEAD} = 1, 1 \leq \text{adj}]$

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

In his approach, $\leq$ means that something is less-than-or-equal-to something. That is, in (9), the adjective *wealthy* and the noun *Republican* specify the upper bound on their HEAD values. That is, their HEAD value $1$ is a supertype of *adjective* and *noun*. Therefore, the HEAD value of the coordinate mother phrase will be neutralized to the nonmaximal value *nominal*. Note that this proposal makes it possible to maintain the strong parallelism of coordination.

Sag’s approach also provides a neat solution to the complicated problem given in (10) (cf. Pullum and Zwicky 1986).

(10) a. I certainly will, and you already have, {clarify / clarified the situation}

    b. I certainly will, and you already have set the record straight with respect to the budget.

The right node raising construction given in (10b) cannot be solved in an approach where the underspecification is not allowed in syntactic domain. Therefore, it shows that the underspecification as proposed in Sag (2003) is inevitable in dealing with syncretic words that are neutral over the relevant aspects. Sag (2003) suggests the following lexical entries for the verbs in (10).
As shown in (11), VFORM value of set is defined to be past participle (pp) or base form (bse) and resolved to the conjunctive type pp&bse. On the other hand, clarify and clarified fix their VFORM values as bse and pp, respectively. As a result, as shown in (11), the verb set satisfies both the selectional restrictions imposed by will and have whereas the verb clarify and clarified do not.

In this paper, I adopt Sag’s approach, which eliminates the requirement that feature structures should be maximally resolved. This is because I believe that the relaxation of the tight feature logic is necessary not only in the lexical domain, but also in the syntax domain. By doing this, I will show that the alternative approach can avoid various problems that are otherwise caused by the framework equipped with the tight feature logic.

4. An underspecification based HPSG approach

In section 2, I showed that natural languages may not in fact coincide with the assumption that feature structures must be fully specified. Thus, I suggests that we need to alleviate the rigidity of the feature system and adopt the underspecification analysis of Sag (2003). In this section, I will show how the alternative approach can explain the natural language phenomena that otherwise may remain unresolved.

4.1 The underspecified case value of who

In section 2, I introduced the mixed case approach by Levine, Hukari and Calcagno (2001). In their approach, the relative pronoun who satisfies any selectional requirement for nominative and accusative. This approach allows relief from the tight feature logic.

(12) a. {Whom / who} will you go there with?
b. With \{whom / *who\} will you go there?

In the approach, the interrogative \textit{who} can appear on the filler position, which is supposed to share the local value with the gap of the object of the stranded preposition. Suppose that the preposition \textit{with} requires that the case value of its object gap be \textit{sacc} in the type hierarchy in (13).

\[(13)\]
\[
\text{s核酸} \\
\text{核酸} \quad \text{sn核酸} \\
\text{acc} \quad \text{nom-核酸} \quad \text{nom}
\]

If the case value of \textit{who} is \textit{nom-核酸} which is a subtype of both \textit{s核酸} and \textit{sn核酸}, it is compatible with \textit{s核酸}. A similar explanation can be applied to the parasitic relatives in (4).

However, the approach suggested by Levine, Hukari and Calcagno (2001) seems to be too rough to cover all the phenomena exhibited by \textit{who}. For instance, as shown in (12b), the interrogative \textit{who} cannot be adjacent to the preposition \textit{with}, but only the accusative interrogative \textit{whom} can. I suggest that the phenomena can be well explained by Sag’s underspecification based account. In this approach, the CASE value of \textit{who} is \textit{nom-核酸} whereas that of \textit{whom} is \textit{核酸}.

\[(14)\]
\[
\text{a. who } [\text{CASE} = [1], [1] \leq \text{nom-核酸}] \\
\text{b. whom } [\text{CASE} = [1], [1] \leq \text{核酸}]
\]

To deal with the distributional differences exhibited in (12), consider the following phrasal structure rules:

\[(15)\]
\[
\text{a. hd-subj-ph: SUBJ} = ( [\text{CASE} = [1], \text{sn核酸} \leq [1]]) \\
\text{b. hd-comp-ph: COMP} = ( [\text{CASE} = [1], [1] \leq \text{核酸}])
\]
Sag (2003) assumes that lexical entries specify the appropriate maximal types. Likewise, the phrase structure rules can be assumed to define the appropriate maximal values. Hence, the phrasal structures are enough to explain the data in (16).

(16) a. Who loves you?
    b. *Whom loves you?

Since the subject should have the CASE value whose supertype is snom, it can have either nom or nom-acc. As a result, who, but not whom, is compatible with this requirement. On the other hand, the structural constraint given in (15b) can be applied differently depending on the situation. When the complement is realized in situ, it is supposed to have acc for its CASE value. When it is not, the CASE value that is percolated onto the top through slash feature becomes the supertype of acc. Therefore, the CASE value of the slashed item becomes compatible with the case values of the lexical entries, who or whom, whose CASE values are also the supertypes of nom_acc and acc, respectively.

4.2 Double nominative constructions in Korean

Kim (2000) suggests that the nominative marker of the sentence initial phrase in Korean double nominative constructions (hereafter DNCs) is solely a focus marker, not the nominative marker. For this argument, Kim (2000) provides various pieces of supporting evidence. First, only the first i/ka marked phrase can be wh-questioned.

(17) a. Nwu-ka apeci-ka kyoswu-i-si-ni?
   who-NOM father-NOM professor-COP-HON-Q
   `(lit.) Who is it whose father is a professor?'
   b. *John-i nwu-ka kyoswu-i-si-ni?
      John-NOM who-NOM professor-COP-HON-Q
      `(lit.) John's `who' is a professor?'

Second, only the initial NP can function as an independent focus phrase in cleft constructions as shown in (18).
John-NOM/GEN computer-NOM broke
‘It is John whose computer is out of order.’

computer-NOM broken-PN man-TOP John-COP-DECL
‘The man whose computer is broke is John.’

John-NOM broken-PN thing-TOP computer-NOM
‘The thing which is broken among John’s stuff is computer.’

According to Kim (2000), the data in (17) and (18) show that the syntactic status of the initial NPs is distinguished from that of general nominative NPs. On the basis of the observations, Kim (2000) proposes that the –i/ka marker of the NPs is a focus marker (see also Kim, Sells and Yang 2007). In Kim’s approach, –i/ka markers allow two different lexical descriptions as in (19):

(19) a. nominative marker  
\[
\begin{array}{c}
\text{marker} \\
\text{MARKING } i/ka \\
\text{CASE nom}
\end{array}
\]  
b. focus marker  
\[
\begin{array}{c}
\text{marker} \\
\text{MARKING } i/ka \\
\text{FOCUS 1}
\end{array}
\]

Leaving the appropriateness of the approach to future studies, let me accept Kim’s account in which the initial phrase of DNCs is actually marked with a focus marker. Then, the dilemma is how the NP preceding the coordinate structure in (20) should be explained.

(20) Ken-i [cha-ka kocangna-ss-ko ton-i up-ta]
Ken-NOM car-NOM break-PST-CONJ money-NOM have.no-DECL
‘Ken’s car was broken, and Ken has no money.’

In Kim (2000)’s account, the initial NP of the Korean DNC in (20) will be specified in a SPR list through head-focus schema. Therefore, in the first conjunct, the unsaturated item is a specifier but in the second conjunct, it is a subject. In addition, the feature structure of the specifier includes the specification of focus
information whereas that of the subject includes the nominative case specification. The tree structure in (21) shows that this type of account cannot survive.

The dilemma is caused by the special status of the initial NP in Korean DNCs. Regarding this issue, the idea in Choi and Lee (2008) provides a solution. According to Choi and Lee (2008), the special status is not limited to the initial NP of the DNCs. Instead, any member of DPES list can appear in the SPR list when the other part of the clause carries important information enough to characterize it. That is, the semantic constraint for the SPR list membership is an aboutness condition (Kang 1998, O. Grady 1991, Hong 1997, Yoon 2004). There is ample evidence supporting this idea. For example, as argued in Choi and Lee (2008), any dependent of a sentence can appear in the raising controller position of ECM constructions as long as it satisfies the aboutness constraint.²

(22) a. na-nun Ken-ul [e apeci-ka kyoswu-la-ko] syangkakhyass-ta
   I-TOP Ken-ACC father-NOM professor-COP-COMP thought-DEC
   'I thought Ken's father was a professor.'

b. na-nun LA-lul [e (mikwuk-yese) hankwuksalam-i cyeil manhi
   I-TOP LA-ACC America-LOC Korean-NOM most a lot

² Choi and Lee (2008) further suggest that the embedded clause of Korean ECM constructions further satisfy the semantic condition for individual level predicates in the sense of Carlson (1977).
santa-ko syangkakhan-ta
live-COMP think-DEC
'I think LA has the most Koreans (in America).'
c. na-nun ecye-lul [e ol eylum cwung nalssi-ka kacang
I-TOP yesterday-ACC this summer during weather-NOM most
tewessta-ko] syangkakhan-ta
be.hot-COMP think-DEC
'I think yesterday was the hottest day this summer.'
d. na-nun Ken-ul [e cheneya-la-ko] syangkakhan-ta
I-TOP John-ACC genius-COP-COMP think-DEC
'I think John is a genius.'

On the other hand, an element that does not satisfy the aboutness condition
cannot appear in the raising controller position as shown in (23) (Choi and Lee
2008).

(23) ??na-nun Chelswu-lul [e swul-ul masikoissta-ko] syangkakhyass-ta
I-TOP Chelswu-ACC alcohol-ACC drink-COMP thought-DEC
'I thought Chelswu drinks.'

In addition, the initial NP in (20) satisfies the aboutness condition in both
conjuncts. That is, the properties, Ken’s father’s job and being a genius, are
important enough to characterize Ken. If this condition is not maintained, the
resulting coordinate structure causes the awkward interpretation as in (24).

(24) ??Ken-i [apeci-ka kyoswu-i-si-ko
Ken-NOM father-NOM professor-COP-HON-CONJ
swul-ul masi-koiss-ta]
alcohol-ACC drink-PROG-DECL
‘Ken’s father is a professor, and Ken is drinking.’

What we observe here is that the distributional behavior of the initial NP in
Korean DNCs coincides with that of the NPs which satisfy the aboutness condition
in other constructions. This observation supports the appropriateness of the lexical
rule in (25) (Choi and Lee 2008).

(25) SPR lexical rule (optional)\(^3\)

\[
\begin{align*}
\text{HEAD} & \quad \text{verb} \\
\text{DEPS} & \quad \ldots, [\text{NP}_i], \ldots \\
\text{INDEX} & \quad i \\
\text{RELS} & \quad \ldots, \left[\text{PRED} \begin{array}{c} \text{aboutness} \\ \text{ARG1} i \\ \text{ARG2} s \end{array}\right], \ldots
\end{align*}
\]

\[\rightarrow \]

\[
\begin{align*}
\text{verb} \\
\text{VAL} \left[\text{SPR} \left(\left[\text{NP}_i\right]\right)\right]
\end{align*}
\]

Now the remaining problem is how to reconcile the conflict between the focus marker and the nominative marker. Building on Kim (2000), I assume that the lexical entries of \(i/ka\) marked entities are organized in a hierarchy in essentially the same as the case types in (13).

(26)

In my approach, the conjuncts of the coordinate structure in (20) have unsaturated specifiers. The feature structures of the specifiers are illustrated in (27).

---

\(^3\) This rule states that any member of dependents can optionally appear in the SPR list as long as it satisfies the \textit{aboutness} condition.
When the feature structures of the two SPR lists undergo unification, the feature structure will be resolved to the feature structure in (27b), which is more specific. A potential problem of this approach is that the feature logic I loosened for the analysis is not the requirement of the maximal resolution, but the principle of totally well-typed feature structures. In this sense, my approach is not totally identical to that of Sag (2003), while following the basic idea of the approach. This type of deviation does not result in any type of chaos of feature logic because the type hierarchy still defines what features are underspecified in a type. Instead, this approach explains the complicated feature mismatch without undermining the generalization and constraints of coordinate structures in Pollard and Sag (2004). Most importantly, the unsaturated VALENCE lists of the conjuncts are not in conflict in my approach and hence share the CATEGORY and NONLOCAL values.

### 4.3 Korean sentential nominalization in coordination

In section 2, I mentioned that the mixed category approach to Korean sentential nominalization in Choi (1999) cannot overcome the category mismatch problem without neglecting the requirement that features are maximally resolved. However, if we ease the tight requirement for the feature structure as we have done in the previous sections, such a problem can be easily avoided. The coordinate structure contains two categories that are in the subsuming relation as revealed in the type hierarchy in (13). That is, the first conjunct in (5) ends with a word with the HEAD value *Verbal* whereas the second conjunct with the HEAD value *VN*. The specific Head feature description for the lexical items is given in (28).
The two feature structures given in (28) undergo unification in the coordinate structure and are resolved to that in (28b). The specific mechanism depends on the unification of feature structures. That is, the lexical item with the feature description in (28a) can be realized as either a word with the HEAD value $V$ or a word with $VN$. This is because the feature description allows its HEAD feature to underspecify the maximal type. Since the head word of the final conjunct has the more specific value $VN$, the whole coordinate structure is resolved to $VN$. The whole coordinate structure now becomes a subtype of nominal so that it combines with an accusative case marker and appears in the object position of the sentence as in (5).

5. Conclusion

In HPSG framework, it has been maintained that the feature structure should be maximally resolved and totally well-typed following the foundational work in King (1989, 1994) and Pollard and Sag (1994). However, this tenet has posed various restrictions that prevent scholars from capturing the generalization of phenomena appropriately and modeling the natural languages faithfully.

Building on the underspecification approach in Sag (2003), I proposed an HPSG approach which lets us avoid various complicated problems not only in coordination but also in various extraction constructions. While easing the tight feature logic originally adopted in HPSG, this approach still maintains the rigorous mechanism of feature logic in that all possibilities of type resolution and feature membership are constrained in the type hierarchy and the lexical entries. It seems to me that this type of modification in feature logic is in accord with the current trend in which the feature regime of HPSG operates in constructions typed in a hierarchy.
References


Sag, I. 2003. Coordination and Underspecification In Jongbok Kim and Stephen Wechsler

Incheol Choi
Department of English Education
Kyungpook National University
80 Dahak-ro Buk-gu
Daegu, 702-701, Korea
Tel.: 82-53-950-5831
E-mail: incheol@knu.ac.kr

Received: 2011. 07. 15
Revised: 2011. 08. 08
Accepted: 2011. 08. 12