Semantics of Complex Sentences in Japanese

Hiroshi Nakagawa Shin-ichiro Nishizawa

Dept. of Electronics and Computer Engineering, Yokohama National University
e-mail: { nakagawa, shin }@naklab.dnj.ynu.ac.jp

Abstract

The important part of semantics of complex sentence is captured as relations among semantic roles in subordinate and main clause respectively. However if there can be relations between every pair of semantic roles, the amount of computation to identify the relations that hold in the given sentence is extremely large. In this paper, for semantics of Japanese complex sentence, we introduce new pragmatic roles called *observer* and *motivated* respectively to bridge semantic roles of subordinate and those of main clauses. By these new roles constraints on the relations among semantic/pragmatic roles are known to be almost local within subordinate or main clause. In other words, as for the semantics of the whole complex sentence, the only role we should deal with is a *motivated*.

1 Introduction

Our aim is to formalize constraints that are needed to develop a parser based on unification grammar (called "UG" henceforth) so that our parser can deal with variety of types of sentences in Japanese. However just parsing syntactically is not enough for natural language understanding. One important and necessary task to be done, when a parser processes a discourse in Japanese, is the so called zero anaphora resolution. All of syntactic, semantic, and pragmatic constraints are to be involved to resolve zero anaphora. Of course, some of omitted pronouns are syntactically resolved. For instance, VP with suffix *te* is not regarded as a clause but a conjunct VP. Therefore the subject of the VP with *te*, which is possibly omitted from surface, should corefer with the subject of the sentence. One example is

(1) Hanako -wa / -TOPIC ϕ_{1subj} samuku-te / feel cold
ϕ_{2subj} mado / window - o / - ACC sime-ta. / closed.
'Hanako felt cold and closed the window.'

where both of zero subjects ϕ_{1subj} and ϕ_{2subj} [1] refer to the sentential topic *Hanako* [2] . In this example, one of the possible accounts for this interpretation is the following. Zero subject of *-te* phrase is [-|- anaphoric, -|- pronominal] or PRO in GB term [Sells 85]. As the result, ϕ_{1subj} is controlled by the subject ϕ_{2subj} of the main VP, which is also zero subject. ϕ_{2subj} is, in GB term, [- anaphoric, -|- pronominal] or pro. The sentential topic Hanako is the only possible antecedent of this zero subject in this example. However, in complex sentences, things are quite different. Consider the following sentence.

(2) Hanako - wa / - TOPIC [ϕ_{1subj} / [samu / feeling cold
-gat-ta / behaved like node] / because]
ϕ_{2subj} mado - o / window - ACC sime-te / close yat-ta. / gave.
1. 'Since Hanako behaved like feeling cold, I closed the window.'
2. 'Since I behaved like feeling cold, Hanako closed the window.'

If contextually we can take only Hanako and the speaker of this sentence as candidates of antecedent of ϕ_{1subj} or ϕ_{2subj}, intuitively the following two interpretations are equally likely.

a. ϕ_{1subj} = Hanako, ϕ_{2subj} = speaker

b. ϕ_{1subj} = speaker, ϕ_{2subj} = Hanako

Therefore ϕ_{1subj} and ϕ_{2subj} are both pro. In fact this fact is well known among Japanese linguists, i.e. [Sells 85, Takubo 87]. As a result, zero anaphora resolution of complex sentence is not only to be done syntactically, but also to be done pragmatically and/or semantically. One of the promising candidate for this is the centering theory [Brennan et al 87, Walker 90]. To apply the centering theory that is originally for a sequence of sentences, namely *discourse*, we regard the subordinate clause and the main clause as a segment of discourse respectively. Moreover Hanako who is marked by 'wa' is regarded as the topic for these two clauses. Then, the topic

[1] Henceforth, $\phi_{\$\$\$}$ means zero \$\$\$.., where \$\$\$.. is either grammatical, semantic or pragmatic role. For instance, ϕ_{subj} means zero subject,ϕ_{agt} means zero *agent*, ϕ_{exp} means zero *experiencer*, and so forth.

[2] 'Hanako' is a typical girl's name.

Hanako is the strongest candidate for the backward center of the subordinate clause. Therefore the backward center of the subordinate clause is Hanako, and consequently zero subject ϕ_{1subj} refers to Hanako. By the same way as the subordinate clause case is dealt with, the zero subject of the main clause ϕ_{2subj} is known to refer to Hanako, too. This result is neither interpretation **a** nor **b** shown above. Another candidate is the property sharing thoery [Kameyama 88]. In her theory, since the both of zero subjects share the subjecthood, both of them finally are known to refer to Hanako that is the topic for both of these clauses. Therefore the property sharing theory also fails to account for the intuitive interpretations.

Then we shift our attention to more microscopic one, in which ,roughly speaking, the important part of semantics of complex sentence is formalized as relations among semantic roles that appear in the main clause or the subordinate clause. At the first glance, the constraints about these relations are not local in terms of main or subordinate clauses. In other words, semantic roles that appear in subordinate clause and semantic roles that appear in the main clause seem to be directly constrained by the constraints of complex sentence. However, looking more carefully, we find that the constraints of subordinate clause and the constraints of main clause are represented as local constraints by introducing the new notion of *motivated* which is characterized as a person who has enough reason to act as the main clause describes. More precisely, *motivated* is one of the pragmatic roles that appear in a subordinate clause, and the constraints in subordinate clause are stated as identity relations between *motivated* and other semantic/pragmatic roles appearing in subordinate clause. Therefore these constraints are local in subordinate clause. The constraints in main clause are stated as identity relations between *motivated* which comes from subordinate clause, and other semantic roles appearing in main clause. Therefore in understanding the main clause we don't have to be care about semantic/pragmatic roles in subordinate clause other than a *motivated*. In this sense, the constraints in the main clause can be treated as almost local constraints of the main clause.

The next question is how to represent the semantics of complex sentence in feature structure(called FS henceforth). For this, we should write down the constraints about these relations among semantic/pragmatic roles in a feature structure formalism. Due to the space limitation, in this paper we mainly pursue the constraints about semantic feature structures.

2 Hierarchical Structure of Complex Sentence

We pay our attention to the general structure of Japanese utterance which is helpful to represent semantics of complex sentence. Several Japanese linguists have already proposed the general structure of Japanese utterances [Mikami 53, Minami 74, Takubo 87, Gunji 89]. Mikami categorized clauses into three classes, namely 'open', 'semi-open' and 'closed.' This categorization indicates how freely the content of clause interacts with the outside of clause. For instance, they are categorized by the degree of possibilities of coreference between zero pronouns inside the subordinate clause and nominal or topic that appear in the main clause. Following Mikami's idea, Minami proposed four levels, namely level A, B, C and D which correspond roughly to VP, proposition, sentence without communication mood and utterance which takes into account a hearer, respectively. [Takubo 87] divided level A into two levels. One of them corresponds to VP, the other corresponds to VP + a certain kind of subject which is called "objective subject." Gunji proposed the more detailed structure, in which starting from predicate, say, verb and adjective, objects, voice, subject, aspect, tense, modality, topic and mood are or might be sequentially added to make an informationally more fulfilled sentence component. Finally, it ends up with an utterance. In Gunji's structure, some node can have more than two daughter nodes to make more complex sentence. Following them, the structure of the so called (cluase level) complex sentence is the following shown in Fig.1.

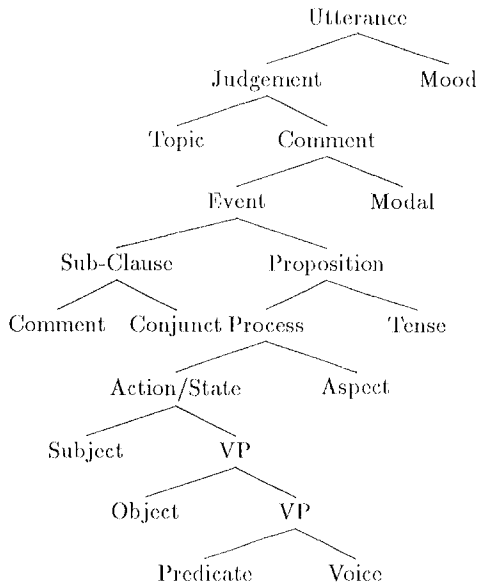

Figure.1: The hierarchical structure of Japanese utterances

In Fig.1 , Sub-Clause and Conjunct mean subordinate clause and conjunctive particle respectively. Note that Fig.1 represents not only the hierarchical structure but also the word order of a complex sentence in Japanese. The structure is almost the same as Gunji's structure except for explicitly showing complex proposition, subordinate-clause

and conjunctive-particle that are newly added to deal with complex sentences. Note that 'Comment' appearing in 'Sub-Clause' has the same structure as 'Comment' appearing just below 'Judgement'. That is to say, 'Comment' is recursively defined. However, in practice, the more the level of depth of recursively appearing 'Comment' is, the less comprehensible the sentence is.

3 Subordinate Clause

In this section, at first we show the predicate categories used in the subordinate clauses that we deal with in this paper, in Table.1. In each category of 2,3,4,5 and 6, exists there a person who is affected by the situation described by the subordinate clause. On the contrary, in category 1, there is not necessarily an explicit affected person. In our theory, this affected person plays a key role for semantics of complex sentence. As the result, in general we cannot derive a useful result for category 1 in our theory. Therefore we don't deal with category 1 in this paper.

At this moment, we should explain the nature of the so called subjective predicate mentioned in Table.1. In short a subjective predicate describes the *experiencer*'s inner state which can exclusively be known by the *experiencer* him/herself.

Next we focus on verbal suffix *garu*. Firstly we show *garu*'s syntax. *Garu* is the present form and its root form is *gar*. Therefore inflections are as follows: *gar-re,gar-i*, etc. In addition, *garu* has an allophonic root form *gat* and, *gat-ta*(past-form), *gat-teiru*(progressive-form) and so on are derived from *gat*. Some of these forms will appear in our examples. Next we talk about the semantics of *garu*. *Garu* roughly means "show a sign of" or "behave like ..ing"[Ohye 75]. Also in [Palmer 86] its semantics is informally explained, however our proposal is to formalize *garu*'s semantics in UG or more generally in computational linguistics. For this, first of all, we introduce a new pragmatic role called *observer*.

Definition 1 (Observer) *Observer is a person who directly observes or is indirectly informed the situation described by the proposition part. Therefore an observer has a certain evidence to be convinced that that situation actually happens.*

1	non-subjective predicate
2	subjective verb
3	subjective adjective without verbal suffix *garu*.
4	subjective adjective with verbal suffix *garu*.
5	*verb* + *ta-garu* (behave as s/he wants to "*verb*")
6	transitive passive and intransitive passive (adversity passive).

Table 1: Predicate Categories

Although this notion of *observer* shares a large part with PIVOT of [Iida-Sells 88], our notion of *observer* is introduced only by *garu*. Therefore it is much narrower notion. As you will see later, this newly introduced role is playing a key role which bridges semantic roles of subordinate clause to semantic roles of main clause.

As for an *observer* introduced by *garu*, one of the widely known consequence about the nature of subjective predicate is the following. In a sentence, if a subjective adjective is used without being followed by a verbal suffix *garu*, the *experiencer* of the subjective adjective should be the speaker of the sentence.

The next thing we should do about a newly introduced notion of *observer* is to make clear the way to deal with it in FS. First of all, in our FS, a semantic content:SEM is basically a soa (state of affair) form of situation semantics. However we use semantic role like "*agent*", "*patient*", "*experiencer*", and so on, as argument roles of soa. Since an *observer* observes the situation which is characterized by a soa, if we know that there exists an *observer*, the observed soa is embedded in observing situation, which, in turn, is embedded in the whole semantic content. In this sense, the observed soa's argument role is *observed*. But as far as we have no confusion, we omit role name 'observed' henceforth. A typical schema of SEM of FS of this type is the following. Note that we use *garu* as a value of the relation feature meant by 'rel.' The English gross of this relation *garu* is 'observe.'

$$(3)\ \text{SEM} = \begin{bmatrix} \text{rel: } garu \\ \text{observer: } \boxed{\text{o}} \\ \text{soa: } \begin{bmatrix} \text{rel:R} \\ \text{agent: } \boxed{\text{a}} \\ \text{experiencer: } \boxed{\text{e}} \\ \text{patient: } \boxed{\text{p}} \\ \ldots. \end{bmatrix} \end{bmatrix}$$

Now we explain the semantics of clause which consists of subjective adjective with *garu* or *ta-garu*, that are in categories 4 and 5. These categories' forms are "ϕ_{exp} P-garu" or its past form "ϕ_{exp} P-gat-ta", where P is a subjective adjective (category 4 in Table.1) or is a verb followed by *ta-gar* (category 5 in Table.1), and ϕ_{exp} is the *experiencer* of P which is possibly zero. In these categories, there exist observers who are not the *experiencer* of P, and observe that experience. The SEM feature of "ϕ_{exp} P-garu/gat-ta" is the following.

$$(4)\ \begin{bmatrix} \text{rel:}garu \\ \text{observer: } \boxed{\text{o}} \text{ where } \boxed{\text{o}} \neq \boxed{\text{e}} \\ \text{soa: } \begin{bmatrix} \text{rel:P} \\ \text{exp: } \boxed{\text{e}} \end{bmatrix} \end{bmatrix}$$

where " $\neq$ " means "not token identical."

In our FS, constraints for tokens like $\boxed{\text{o}}$ are written with "where" as shown in this FS. Since constraint satisfaction method in UG has been and is developed by many researchers recently i.e. [Tsuda 91],

our theory will be able to be implemented in systems like theirs.

If the sentence finishes just after "garu/gat-ta", the important points are 1) an introduced *observer* is the speaker, and consequently 2) the *experiencer* cannot be the speaker. If a clause with "garu/gat-ta"is a subordinate clause, the *experiencer* cannot be identified with a semantic role corresponding to the subject of main clause or higher clause.

As for category 2, subjective verbs like "kurusimu"(feel sick) and "kanasimu"(feel sadness) that describe subjective and/or emotional experience in verb form, are used. Like the case of *garu*, an *observer* who observers the experience can be introduced. However this observer is not obligatory. Therefore unlike the "garu/gat-ta" case, the *experiencer* also can be an obligatory semantic role of higher clause as well as the speaker.

4 Complex Sentence

4.1 Feature Structure

According to the hierachical structure of Japanese sentence shown in Fig.1 , the essential part of hierarchical structure of the following sentence (5) is shown in Fig.2 . In this figure, the structure just below each proposition is replaced with the corresponding parts of sentence.

(5) [ϕ_{exp} samu -gat-ta
[feel cold behaved like
node] ϕ_{agt} mado -o sime-ta.
because] window -ACC closed.
'Since ϕ_{exp} behaved like feeling cold, ϕ_{agt} closed the window.'

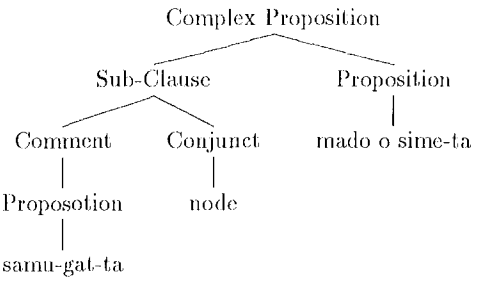

Figure.2 : Hierarchical structure of (5)

Basically the embedding structure of FS corresponds to the hierarchy shown in the hierarchical structure Fig.1 . To grasp the image of the relation between a hierarchical structure and the corresponding FS, we show an example of FS of the above complex sentence (5) analyzed based on this hierarchical structure in the following. This FS is the result of the unification between the FSs of subordinate clause and main clause, where the contents of syntactic feature HEAD , namely [syn] is omitted.

$$\left[\begin{array}{l}\text{MORPH: 'samu-gat-ta node,mado o sime-ta'}\\ \text{HEAD: } \boxed{\text{syn}}\\ \text{SEM: } \boxed{\text{matrix-sem}} \wedge \boxed{\text{sub-sem}} \wedge \boxed{\text{o}} \neq \boxed{\text{c}}\end{array}\right]$$

$$\boxed{\text{matrix-sem}} = \left[\begin{array}{l}\text{rel: } \textit{sime}\\ \text{agent: } \boxed{\text{o}}\\ \text{object:window}\\ \text{tense:past}\end{array}\right]$$

$$\boxed{\text{sub-sem}} = \left[\begin{array}{l}\text{rel: } \textit{node}\\ \text{motivated: } \boxed{\text{o}}\\ \text{soa: } \left[\begin{array}{l}\text{rel: } \textit{garu}\\ \text{observer: } \boxed{\text{o}}\\ \text{soa: } \left[\begin{array}{l}\text{rel: } \textit{samu-i}\\ \text{experiencer: } \boxed{\text{c}}\\ \text{tense:past}\end{array}\right]\end{array}\right]\end{array}\right]$$

where English grosses of relation name is the following: *sime*:'close', *node*:'because', *samu-i*:'feel cold'.

The key point of the semantics of complex sentence is the role **motivated** that appears in [sub-sem] which corresponds to the content of the subordinate clause. The role *motivated* is the link between the content of subordinate clause and the main clause. Semantically *motivated* is characterized as the following.

Definition 2 (Motivated) *Motivated is a person who is affected by the situation described by the subordinate clause deeply enough to feel or act as the main clause describes.*

The important and indispensable part of semantics of complex sentence is, roughly speaking, the relation between a subordinate clause and the main clause. But if you look more closely, this relation is actually the relations among semantic/pragmatic roles appearing in the subordinate clause and those appearing in the main clause. The newly introduced role of *motivated* gives the most important clue for this relation. Therefore, in the rest of this paper, our effort will be concentrated into whom a *motivated* refers to. More precisely, in FS, our main concerns are which semantic role in the SEM of subordinate clause the *motivated* can or cannot be unified with, and which semantic role in the SEM of main clause the *motivated* can or cannot be unified with.

4.2 Constraints

In this subsection, we propose the constraints on complex sentence. For this, at first we categorize the relations between subordinate clause and main clause based on their semantics. They are divided up to many types of complex sentence. We show the most important and typical types in Table.2, where SC and MC mean 'subordinate clause' and 'main clause' respectively. In this table, the first column is for a name of sentence type, the second column indicates a rough meaning of the relation between

type	outline meaning of complex sentence	Japanese conjuncts
1	SC causes MC	*node, kara*
2	Although SC, MC	*noni, ga, keredomo, temo, i-te, i-tutu, i-nagara*
3	If SC then MC	*to, nara, tara, reba*
4	When/after/before etc SC, MC	*toki, ato, mae,* etc.

Table 2: Clause Adjuncts

subordinate clause SC and main clause MC of complex sentence, and the third column shows Japanese conjunctive particles used to represent a type of complex sentence in the same row.

Three VP adjuncts, *te, tutu,* and *nagara*, are usually used to express events ocurring simultaneously. However, if they are used with aspectual suffix *i* which means perfective, for instance *i-nagara*, they are regarded as clause conjuncts and are to be interpreted as 'although'[Minami 74]. We don't deal with type 4, because a temporal adverbial clause just describes an event that occurs before, simultaneously or after another event which is described by the main clause. Therefore generally we don't expect essential information for relations among semantic roles appearing in adverbial or main clause from this type of sentence.

Now we focus on type 1,2 and 3, where a *motivated* plays the key role in the constraints. In Table.3 we show the constraints that say which semantic/pragmatic role of subordinate clause can be a *motivated*. Table.4 shows which semantic role of main clause can be unified with the *motivated*. In these tables, the first column of the first row is for constraint names, the second column shows a set of sentence types for which the constraints shown in the second row apply. The third column of Table.3 shows predicate patterns of subordinate clause, and the third column of Table.4 shows semantic categories of predicate of main clause. For them, constraints written in the second row apply. Note that all of these constraints in Table.3 are local in a subordinate clause, because both sides of $=$ of constraints are roles of subordinate clause. In case of subjective adjective without *garu*, the constraint '*motivated* $=$ *experiencer*' holds also for type 1 except for the case where directionally auxiliary verb "yaru(give)", "kureru(be given)" are used. Analysis for these cases is one of our future problem.

As for Table.4, state* is a state except for the case that there exists a third party who is a *motivated* puts the *experiencer* into that state. For instance, the *experiencer* is permitted to do something by the *motivated*. Since in this kind of case things are quite complicated, we omit it here because of the limited space. Constraints in Table.4 are also local in a main clause because every semantic role that appears in the righthand side of the constraints is defined within

name	type	predicate
	constraint	
S1	1,3	subjective adjective + *garu* verb + *ta* + *garu*
	motivated = *observer*	
S2	2	subjective adjective + *garu* verb + *ta* + *garu*
	1,2,3	subjective verb
	motivated = *observer* $\vee$ *experiencer*	
S3	1	subjective adjective (without *garu*)
	motivated = *experiencer*	
S4	2,3	subjective adjective (without *garu*)
	motivated = *experiencer* $\vee$ *observer*	
S5	1,2,3	intransitive passive
	motivated = *affected*	
S6	1,2,3	transitive passive
	motivated = *affected* :if *affected* exists, = *patient* : otherwise	

where 'name' means a name of each constraint.

Table 3: Constraints in Subordinate Clause

name	type	predicate category
	constraints	
M1	1,2,3	action
	sub-clause:SEM:*motivated* = *agent*	
M2	1,2,3	state*
	sub-clause:SEM:*motivated* = *experiencer*	

Table 4: Constraints in Main Clause

the main clause. Needless to say, the influence from a subordinate clause comes only via role *motivated*.

In the rest of this section we show the examples that exemplify these constraints. [3]

First, we take (5) of type 1. The constraints to be applied are S1 and M1 as you know from the contents of subordinate and main clause. By combination of S1 and M1, zero *agent* of main clause:ϕ_{agt} is the *observer* of the situation described by the subordinate clause, where ϕ_{exp} behaved like feeling cold. This interpretation coincides with native's intuition.

Look at the following pair of example.

(6) [ϕ_{exp} kurusi -gat-ta noni]
[feel bad behaved but]
kekkyoku ϕ_{agt}
at last
kusuri - o nom anakat-ta.
medicine - ACC drink not-PAST.
'Although ϕ_{exp} behaved like feeling bad, ϕ_{agt} didn't take a medicine at last.'

(7) [ϕ_{exp} nokori ta -gat-ta
[stay want behaved like
noni] kekkyoku ϕ_{agt} oi dasi-ta
but] finally forced out.

[3]The examples shown below are a tip of iceberg we actually analyzed, of course. We gather the data about native's intuitive interpretation from more than twenty natives around authors.

'Although ϕ_{exp} wanted to stay, ϕ_{agt} finally forced him out.'

In both of (6) and (7), the *motivated*s of subordinate clause are constrained by S2, namely *motivated*s can be either ϕ_{exp} or the *observer* of subordinate clause. Constraint M1 says that in both cases, ϕ_{agt} is unified with the *motivated*. Intuitively in (6), ϕ_{agt} is ϕ_{exp}. On the other hand in (7), ϕ_{agt} is the *observer*. Both of these interpretations comply with constraints S2, and M1.

(8) [ϕ_{1exp} atui node]
[be hot because]
ϕ_{2exp} komaru.
be in trouble.
'Since it is hot, I am in trouble.'

Intuitively ϕ_{1exp} corefer with ϕ_{2exp}. This interpretation is expected by constraint S3 and M2 that apply in this case. As you know from these examples, our constraints are not strong enough to identify the antecedent of ϕ_{agt} uniquely, but makes safe interpretations. Moreover disambiguation done by these constraints is useful for further inference that will be done with commonsense knowledge or with a special vocabulary like 'kekkyoku(finally)' used in (7).

In case of S5, namely intransitive passive or adversity passive, it is well known, i.e. [Gunji 87] that there exists a person who is affected by the situation described by the passive sentence. An example sentence is the following.

(9) [ϕ_{affect} tuma - ni sin -are
[wife be dead -PASSIVE
-ta noni]
-PAST but]
ϕ_{exp} kanasimi - mo -si nai.
show sadness not.
'Although his wife had gone, ϕ_{exp} doesn't show a bit of sadness.

The semantic role of this affected person , in (9) zero role:ϕ_{affect} whose wife was dead, is an *affected*. The intuitive interpretation that $\phi_{exp} = \phi_{affect}(= motivated)$, is expected by our constraints: S5 of Table.3 and M1 of Table.4. On the contrary, in case of S6, namely transitive passive, generally we don't have an *affected*. However in some context, a transitive passive form may require the role *affected* which is inherent to adversity passive. For instance,

(10) ϕ_{affect} saihu - ga nusum
wallet - SUBJ steal
-are -ta
-PASSIVE -PAST
'ϕ_{affect}'s wallet was stolen.'

In this case, a person whose wallet was stolen is not explicit but regarded as an *affected*. Another case having an *affected* is that a relational noun is the subject of transitive passive. Then a person who is in the relation expressed by the relational noun is thought to be affected by that situation ,too. Here we take 'mother', 'father', 'daughter', 'son', 'supervisor', and so forth as a relational noun. A couple of example sentences are the following.

(11) [kobun -ga yar -are
[henchman -SUBJ attack -PASSIVE
-ta node]
-PAST because]
ϕ_{agt} sikaesi - ni it-ta.
retaliate go- PAST
'Since his henchman was attacked, the boss retaliated.'

(12) [kobun -ga yar -are
[henchman -SUBJ attack -PASSIVE
-ta noni] ϕ_{agt} te-o komancite-iru.
-PAST but] did nothing.
'Although his henchman was attacked, the boss didn't retaliate.'

ϕ_{agt} who retaliated (11) (or didn't retaliate (12)) has a certain relation between the henchman who had been attacked. For instance, ϕ_{agt} may be the boss of that henchman. In (11), since constraint S6 of Table.3 and M1 of Table.4 apply, ϕ_{agt} is an *affected* of attacking event described in the subordinate clause. This interpretation coincides with native's intuition.

In sum, with these constraints, a constraint satisfaction process in UG based parsing can be done locally and consequently very efficiently. In other words, primarily a constraint satisfaction process of a subordinate clause can be done within the analysis of subordinate clause, and that of the main clause can be done within it except for using *motivated* whose value has already been constrained in the subordinate clause.

5 Related Works and Conclusions

One of the relevant researches to ours is JPSG that has been developed by Gunji[Gunji 87, Gunji 89] and is further studied by the ICOT working group. Our focus is a more pragmatics oriented one than JPSG is. Many Japanese linguists have already done the enormous amount of basic observations and proposed linguistic theories about the phenomena we deal with in this paper [Mikami 53, Kuno 73, Kuno 78, Ohye 75, Minami 74, Takubo 87, Teramura 84, Teramura 90, Saito 92]. Of course our research is based on their works and observations. In [Ohye 75], it is said that if *garu* is used in a subordinate clause, the subject of the main clause is not the *experiencer* of the subordinate clause. In [Saito 92], she says that 1) a cognizer that corresponds to our *observer* is introduced if *garu* is used, and 2) if an *observer* is introduced in the subordinate clause, the mentally responsible person appearing in the main clause is identical with the *observer*. In linguistic phenomena, these observations are similar to the

constraint we propose here. So what is new? The answer is that: 1) We explicitly state the semantics of complex sentence as the relations among semantic roles. Namely, since we use semantic/pragmatic roles instead of grammatical roles in constraints, our constraints can account for zero anaphora in a sentence where the main clause is passive where an *agent* or an *experiencer* is not necessarily the subject, like the following example.

(13)

Taro	-wa	[gakkou e	iku-no	-wo
	-Topic	[to school	go-NOM	-ACC
iya	-gat-ta	node]		
hate	behaved like	because]		
ϕ_{agt}, ϕ_{pat}	okor	-are	-ta.	
	scold	-PASSIVE	-PAST	

'Since Taro behaved like hating to go to school, he was scolded.'

where the intuitive reading is the following: ϕ_{pat}, that is zero subject, refers to Taro, and ϕ_{agt}, that is not the zero subject, refers to Taro's parents who are the *observer* and *motivated* of the subordinate clause. 2) We formalize this theory in UG formalism, even though the details are omitted due to the space limitation. 3) We find that the constraints of complex sentences are actually local ones. This localization of constraint was found by introducing new pragmatic roles *observer* and *motivated*, and is extremely important for efficiency of UG based parsing. This localization also makes the proposed constraints be compositional ones, because in the case of deeply embedded complex sentence to identify the referent of each *motivated* that bridges between a subordinate clause and its main clause, the constraints we proposed are resolved with computation confined within each clause.

Analysis of case in which a directional auxiliary verb i.e. 'yaru','kureru' is used is left as the future problem. Finally, we implemented a Japanese language understanding system based on the theory we state in this paper, but due to the space limitation we will report the detail of implementation in other place in the near future.

References

[Brennan et al 87] Brennan, S., M. Walker Friedman and C.Pollard (1987). A Centering Approach to Pronouns. 25th Annual Meeting of ACL, pp.155-162

[Gunji 87] Gunji, T.(1987). Japanese Phrase Structure Grammar. Reidel, Dordrecht

[Gunji 89] Gunji,T. (1989). Relevance of the Formalization of Phrase Structure Grammar to Mechanical Language Processing. Report of Tokutei-Kenkyu, Ministry of Education and Academy

[Iida-Sells 88] Iida,M. and P.Sells(1988). Discourse Factors in the Binding of zibun. in Japanese Syntax (ed. W.Poser) CSLI, Stanford

[Kameyama 88] Kameyama, M. (1988). Japanese Zero Pronominal Binding: Where Syntax and Discourse Meet. in Japanese Syntax (ed. W.Poser) CSLI, Stanford

[Katagiri 91] Katagiri,Y. (1991). Perspectivity and Japanese Reflexive 'zibun'. in CSLI Lecture Notes No.26, Situation Theory and its Applications Vol.2, J.Barwise et al eds. pp.425-447

[Kuno 73] Kuno, S. (1973). The structure of the Japanese Language. Cambridge, MIT Press

[Kuno 78] Kuno,S.(1978). Danwa no Bunpou. Taishukan, Tokyo

[Ohye 75] Ohye,S.(1975). Nitieigo no Hikakukenkyu. Taishukan, Tokyo

[Saito 92] Saito,R.(1992). Shinjou Jutugo no Goyouronteki Bunseki (Pragmatic Analysis about Psychological Predicates). Nihongogaku, Vol.11, No.6, pp.110-116

[Mikami 53] Mikami,A.(1953). Gendai Gohou Josetu. Kuroshio-Shuppan, Tokyo

[Minami 74] Minami,F.(1974). Gendai Nihongo no Kouzou. Taishukan, Tokyo

[Palmer 86] Palmer, F.R.,(1986). Mood and Modality. Cambridge University Press,Cambridge

[Sells 85] Sells, P. (1985). Lectures on Contemporary Syntactic Theories. CSLI Stanford

[Takubo 87] Takubo,Y. (1987). Tougokouzou to Bunmyakujouhou (Syntactic Structure and Contextual Information. Nihongogaku 1987-5,Meiji-shoin,Tokyo

[Teramura 84] Teramura,H.(1984). Nihongo no sintakusu to imi II 'Japanese Syntax and Semantics II'. Kuroshio-Shuppan,Tokyo

[Teramura 90] Teramura,H.(1990). Nihongo no sintakusu to imi III 'Japanese Syntax and Semantics III'. Kuroshio-Shuppan,Tokyo

[Tsuda 91] Tsuda,H.,Hasida,K.,Sirai,H. (1989). JPSG Parser on Constraint Logic Programming. 4th ACL European Chapter

[Walker 90] Walker,M.,M. Iida and S. Cote(1990). Centering in Japanese Discourse. COLING'90