

Corpus-Based Analysis of Japanese Relative Clause Constructions

Takeshi Abekawa¹ and Manabu Okumura²

¹ Interdisciplinary Graduate School of Science and Engineering,
Tokyo Institute of Technology, Japan
`abekawa@lr.pi.titech.ac.jp`

² Precision and Intelligence Laboratory,
Tokyo Institute of Technology, Japan
`oku@pi.titech.ac.jp`

Abstract. Japanese relative clause constructions (RCC's) are defined as being the NP's of structure 'S NP', noting the lack of a relative pronoun or any other explicit form of noun-clause demarcation. Japanese relative clause modification should be classified into at least two major semantic types: case-slot gapping and head restrictive. However, these types for relative clause modification cannot apparently be distinguished. In this paper we propose a method of identifying a RCC's type with a machine learning technique. The features used in our approach are not only representing RCC's characteristics, but also automatically obtained from large corpora. The results we obtained from evaluation revealed that our method outperformed the traditional case frame-based method, and the features that we presented were effective in identifying RCC's types.

1 Introduction

Japanese relative clause constructions (RCC's) are defined as being the NP's of structure 'S NP', noting the lack of a relative pronoun or any other explicit form of noun-clause demarcation[1]. Japanese relative clause constructions should be classified into at least two major semantic types: case-slot gapping and head restrictive. However, these types for relative clause constructions cannot apparently be distinguished.

Given the types of Japanese relative clause constructions and a corpus of Japanese relative clause construction instances, we present a machine learning based approach to classifying RCC's. We present a set of lexical and semantic features that characterize RCC's, and integrate them as a classifier to determine RCC types. We use decision tree learning as the machine learning algorithm.

Distinguishing case-slot gapping and head restrictive relative clauses, or resolving the semantic relationship between the relative clause and its head noun has several application domains, such as machine translation from Japanese[5]. It also has a place in text understanding tasks, such as splitting a long sentence into multiple shorter sentences, and removing less important clauses to shorten a sentence[6].

Previously, relative clauses had been analyzed with rule-based methods that utilized case frames[5,2]. Using hand-crafted rules and knowledge creates several problems: the high cost of constructing them, and lower scalability and coverage.

Recently, due to the availability of very large corpora, corpus-based and machine learning-based approaches have been actively investigated[7]. Cooccurrence information between nouns and verbs can be calculated from the syntactically parsed corpus, and this information can be used preferentially instead of hand-crafted case frames to determine whether a noun can be the filler of a case-slot of a verb[7,11].

However, merely using the cooccurrence information between nouns and verbs instead of case frames cannot provide a good solution to the analysis of Japanese relative clauses. Clauses with high occurrence probability of the main verb and the head noun can sometimes be head restrictive. Moreover, just because the head noun can be the filler of a case-slot of the verb does not always mean that the clause as case-slot gapping. We have to rely on several different clues in order to realize accurate classification. Therefore, in this paper we present eight features are effective in classifying case-slot gapping and head restrictive relative clauses. Most of the features can be automatically acquired by statistically analyzing a corpus as explained in section 4.

In section 2 we first describe the nature of Japanese RCC's, and in section 3 we outline previous work on the analysis of Japanese relative clauses. In section 4 we explain the features that we present in this paper, and in section 5 we explain the machine learning-based classifier, which uses the features in section 4. In section 6 we describe the evaluation of the system and discuss the experimental results.

2 Japanese Relative Clause Constructions

Japanese relative clause constructions have the structure '*S NP*', and constitute a noun phrase as a whole. We will term the modifying *S* the "relative clause", the modified *NP* the "head NP", and the overall NP a "relative clause construction" or RCC[2]. Example RCCs are:

- (a) さんまを 焼く 男
 saury grill man
 "the man_i who ϕ_i grills a saury"
- (b) 誰もが 知っている 情報
 everyone know information
 "the information_i which everyone knows ϕ_i "
- (c) さんまを 焼く 匂い
 saury grill smell
 "the smell of saury grilled"

RCC should be classified into at least two major semantic types: case-slot gapping and head restrictive. With case-slot gapping RCC's (also called '**inner**'

relative clauses[14]), the head NP can be considered to have been gapped from a case slot subcategorized by the main verb of the relative clause. Head restrictive RCC's (also called **'outer'** relative clause[14]) occur when the relative clause modifies the head NP. In (a), the head NP “” (man) can be the subject of the main verb of the relative clause, and in (b), the head NP “” (information) can be object of the main verb. These RCC type are 'inner' relative clauses. In (c) the head NP “” (smell) cannot fill the gap in the relative clause, and RCC type is 'outer'.

The inherent difficulty in determining the type of RCC derives from the fact that these two types of RCC are syntactically identical. Even if the relative clause has case-slot gapping, the type of that clause is not always 'inner', because in Japanese the main verb of the relative clause has often zero pronoun. We thus have to disambiguate the individual RCC instances.

3 Related Work

Previous work on analyzing Japanese relative clauses has used case frames as useful information. They have first tried to find the case frame for the main verb of the relative clause and embedded the nouns in the clause into its case-slots. The head noun is then tried to be embedded into the remaining case-slot in the case frame. To determine whether a relative clause instance is 'outer' clause, they have beforehand constructed a dictionary of the nouns that can be modified by 'outer' clause, such as “”(purpose), or “”(opinion). In one approach[5], the instance is determined to be 'outer' clause, if the head noun is included in the dictionary, regardless of the main verb of the relative clause. In another approach[12], the instance is determined to be 'outer', if the head noun cannot be embedded into a case-slot and the head noun is included in the dictionary.

Recently, cooccurrence information between verbs and nouns has been used in analysis. Kawahara and Kurohashi[7] automatically extracted case frames from very large corpora, and used the case frames to analyze Japanese relative clauses. However, they judged the instances as 'outer' clauses, only if case-slot filling did not succeed.

Murata[11] presented a statistical method of classifying whether the relative clause is an 'inner' or an 'outer' clause. However this method cannot correctly classify 'outer' relative clause which had high cooccurrence probability of the main verbs and the head nouns.

4 Feature Set to Classify RCC Type

In this section, we present eight features that can be considered to be effective in classifying 'inner' and 'outer' relative clauses.

1. Degree of possibility where the head noun can be modified by the 'outer' relative clause (degree of 'outerness').

In Japanese, there are two ways of modification between verbs and nouns: nouns modify verbs by filling a case-slot (noun \rightarrow verb), and verbs modify nouns in

Table 1. Comparison of the number of cooccurring verbs

noun	relative clauses		case-slots	
	freq.	verb No.	freq.	verb No.
(intent)	8,732	941	14,216	677
(fact)	5,454	1,448	7,301	754
(preparation)	2,268	428	2,720	74
(people)	6,681	1,367	10,026	1,998
(city)	1,172	449	3,688	857
(television)	2,740	707	30,627	2,228

relative clauses (verb \rightarrow noun). Some pairs of a verb and a noun can cooccur only in RCC, and cannot cooccur by filling a case-slot of the verb. For example, noun “準備” (preparation) and verb “走る” (run) can cooccur with each other as the main verb of a relative clause and its head noun, as in “走る準備” (preparation for running), though the noun cannot fill any case-slots of the verb, as in * “準備が走る” (*preparation runs). For nouns, some verbs only cooccur in relative clauses, and a number of such verbs tend to be modified by ‘outer’ clauses.

Table 1 shows the occurrence frequency of sample nouns and the number of their cooccurring verbs in the relative clauses or in the case-slot relations. For nouns that do not tend to be modified by ‘outer’ clauses, such as “人々” (people), “都市” (city), and “テレビ” (television), the ratio between the frequency and the number of verbs is almost the same between the relative clause and case-slot cases. On the contrary, for nouns that tend to be modified by ‘outer’ clauses, such as “意向” (intent), “事実” (fact), and “用意” (preparation), the number of verbs is much bigger in relation to clause cases, although the frequency is smaller. The reason may be, as previously explained, that some verbs cooccur with the nouns that tend to be modified by the ‘outer’ clause only in relative clause constructions.

Therefore, we can measure the likelihood that the noun will be modified by ‘outer’ relative clauses, by calculating the difference in the frequency distribution of verbs cooccurring in relative clauses against the frequency distribution of verbs cooccurring in the case-slot relation (If the difference is larger, the probability that the noun can be modified by the ‘outer’ relative clause becomes larger).

We calculate the likelihood as $J(P_k(v|n), P_m(v|n))$, the Jensen-Shannon distance between the cooccurrence probability where nouns fill the case-slots of verbs ($P_k(v|n)$) and the cooccurrence probability where verbs cooccur with nouns in relative clauses ($P_m(v|n)$). Given two probability distributions p, q , the Jensen-Shannon distance is defined by the following formula[9]:

$$J(p, q) = \frac{1}{2} \left[D(p \parallel \frac{p+q}{2}) + D(q \parallel \frac{p+q}{2}) \right]. \quad (1)$$

$D(p \parallel q)$ is the Kullback-Leibler distance and defined by the following formula[3]:

$$D(p \parallel q) = \sum_i p_i \log \frac{p_i}{q_i}. \quad (2)$$

Table 2. ‘outerness’ of example nouns

noun	意向	事実	用意	人々	都市	テレビ
	(intent)	(fact)	(preparation)	(people)	(city)	(television)
$J(P_k, P_m)$	0.546	0.360	0.616	0.160	0.155	0.159

We use the Jensen-Shannon distance rather than the Kullback-Leibler distance, because the former is symmetric and has stability in various sizes of probability distribution experimentally. $P_k(v|n)$ and $P_m(v|n)$ are calculated as follows:

$$P_k(v|n) = \frac{f_k(n, v)}{f_k(n)}, \quad (3)$$

$$P_m(v|n) = \frac{f_m(n, v)}{f_m(n)}, \quad (4)$$

where $f_k(n, v)$ is the cooccurrence frequency where noun n fills a case-slots of verb v , and $f_k(n)$ is the frequency of the noun that occurs in the case-slot of verbs. Similarly, $f_m(n, v)$ and $f_m(n)$ are the frequencies for relative clause constructions. Table 2 shows the ‘outerness’ of sample nouns. The values of the nouns that are often modified by ‘outer’ clauses are higher than those of the nouns which tend to be modified by ‘inner’ clauses.

2. Cooccurrence information between head noun and main verb of relative clause.

For a relative clause instance to be an ‘inner’ clause, the head noun has to fill a case-slot of the main verb of the relative clause. Consider the following two examples:

- (a) 共鳴する 音
resonate sound
“the sound_i that ϕ_i resonates”
- (b) 破壊する 音
destruct sound
“the destruction sound”

In (a), “音” (sound) can be the subject (“か” case) of the main verb “共鳴する” (resonate). On the contrary, in (b) “音” cannot fill any case-slots of the main verb “破壊する” (destruct) and can be considered to be modified by the ‘outer’ relative clause. Therefore, if the head noun can fill a case-slot of the main verb, the relation can be more plausibly assessed as ‘inner’.

Whether a noun can fill a case-slot of a verb has been traditionally determined using case frames. However, we use the cooccurrence information between the head noun and the main verb. In this paper, the cooccurrence between nouns and verbs is measured by mutual information. Taking into account the information on case particles, mutual information is calculated with the following formula:

$$I(n, k; v) = \log \frac{p(n, k, v)}{p(n, k)p(v)}, \quad (5)$$

where $p(n, k)$ is the probability that noun n will cooccur with case particle k and $p(n, k, v)$ is the cooccurrence probability for noun n , case particle k and verb v , and $p(v)$ is the occurrence probability for verb v . The following seven case particles were taken into account: (“が”, “を”, “に”, “で”, “へ”, “と”, and “から”). This is because only these case-slots can, in fact, be gapped to the head noun to construct the relative clause.

3. Which case-slots are already filled for main verb by nouns in relative clause.

As previously explained, if the head noun can fill the case-slot of the main verb of the relative clause, the RCC instance can be judged as an ‘inner’ clause. However, if the case-slot that the head noun can fill is already filled by the noun in the relative clause, and hence unavailable for case-slot gapping, the rule cannot be applied. Consider, for example, the following two cases:

- (a) 聞いてきた 話
 hear story
 “the story_i that (someone) heard ϕ_i ”
- (b) 落語を 聞いてきた 話
 Japanese comic story hear story
 “the story that (someone) heard a Japanese comic story”

In (a), since “話” (story) can fill the object (“を” case) case-slot of the main verb “聞く” (hear), the relation can be judged as ‘inner’. However, in (b), since the object (“を” case) case-slot of the main verb “聞く” is already filled by the noun “落語” (Japanese comic story), and “話” cannot fill any case-slot, the instance is judged as ‘outer’.

Taking the situation into account, if a noun in the relative clause fills a case-slot of the main verb, the mutual information for the case-slot is set to a very small value M_{min} .

4. Whether the head noun is modified by modifiers other than the relative clause (other modifier).

Previous work on analyzing Japanese relative clauses has taking into account only the head noun, and has not taking into account modifiers other than the relative clause. Consider the following two examples:

- (a) 彼に 話す 目的
 him talk purpose
 “the purpose that (someone) talk (something) to him”
 “the purpose_i that (someone) talk ϕ_i to him”
- (b) 彼に 話す 旅行の 目的
 him talk trip purpose
 “the purpose of the trip_i that (I) talk ϕ_i to him”

(a) has two interpretations. The first interpretation is that “目的” (purpose) do not fill the remaining case-slots of the main verb “話す” (talk) and can be con-

sidered to be modified by the ‘outer’ relative clause. The second interpretation is that “目的” can be the direct object(“を” case) of the main verb “話す” and can be considered to be modified by the ‘inner’ relative clause. On the contrary, (b) has only the interpretation of ‘inner’.

If the head noun is modified by modifiers other than the relative clause, such as adjectives, compound nouns, and “AB”(B of A), the relative clause type tends to be ‘inner’. The function of ‘outer’ relative clause describes the content of the head noun. If the head noun is modified by a modifier, the relative clause need not describe it. Therefore, the type of relative clause tends to be ‘inner’.

To implement the idea, we define a feature ‘other modifier’. If the head noun is modified by any modifiers other than the relative clause, its value is 1, otherwise, 0³.

5. Whether head noun tends to be modified

As for the nouns which tend to be modified by ‘outer’ relative clauses, the relative clauses describe the content of the head nouns. It is difficult to understand their meaning without any modification. Therefore we calculate the percentage to what degree nouns are modified by any modifier in large corpora. Table 3 shows the percentage for example nouns.

Table 3. Percentage of modification

意向 (intention)	分野 (field)	テレビ (television)	彼 (he)	Average of all nouns
0.983	0.973	0.287	0.155	0.460

The percentages of nouns “意向”(intention) and “分野”(field), which tend to be modified by ‘outer’ relative clause, are close to 1, that is to say, such nouns must have any modification. We consider, the higher this percentage, the higher the possibility that the noun is modified by ‘outer’ relative clause.

6. Percentage where “という” is inserted between relative clause and head nouns

“という” is a function expression that is sometimes inserted between relative clauses and head nouns. Table 4 shows the percentage where “という” cooccurs with example nouns.

Table 4. The percentage of “という” cooccurring with noun

意見 (opinion)	噂 (rumor)	場所 (place)	人々 (people)	Average of all nouns
0.335	0.246	0.007	0.008	0.007

³ In the experiment, we use syntactic annotated corpus. Therefore, other modifier elements are already identified.

The percentages of nouns “意見”(opinion) and “噂”(rumor), which tend to be modified by ‘outer’ relative clause, are higher than the average. We consider, the higher this percentage, the higher possibility that the noun is modified by ‘outer’ relative clause.

7. Whether head noun tends to be modified by past-tensed relative clauses(tense information)

Some nouns tend to be modified by past-tense relative clauses, and others tend to be modified by those in the present tense. Consider, for example, the following two nouns: “計画”(plan) and “記憶”(memory). Both nouns are considered to imply the concept of time (future or past)⁴.

Table 5. Tense of main verb and distribution of inner/outer

tense	計画 (plan)		記憶 (memory)	
	inner	outer	inner	outer
present	6	89	12	0
past	5	0	5	83

For each of the two nouns “計画”(plan) and “記憶”(memory), we examined 100 relative clause instances that had the noun as the head noun (Table 5). If the head noun implies the concept of time, the tense of the main verb of the relative clause tends to coincide with this concept. Furthermore, note that the tense of the main verb of ‘outer’ relative clauses is the same as the time concept of the head noun. From this, if the noun tends to be modified by a specific-tense relative clause, the relative clause tends to be ‘outer’, and if the tense of the main verb contradicts the time concept of the head noun (tense of frequently modified relative clauses), the relative clause should be determined as ‘inner’.

To implement this idea, we first calculated deviations in the distribution of tense for the relative clauses. The percentage of past-tense main verbs in all relative clauses, R_{past} , and the average for all the nouns were calculated. Table 6 shows the results for sample nouns.

Table 6. Percentage of past-tense main verbs

計画 (plan)	記憶 (memory)	場所 (place)	人々 (people)	Average of all nouns
0.032	0.958	0.333	0.422	0.322

For a head noun which does not imply the concept of time (“場所”(place) and “人々”(people)), the percentage is near average. On the contrary, “計画”(plan) and “記憶”(memory) which imply the concept of time have an extreme value.

⁴ In Japanese, there are just two tense surface markers: present and past. Therefore, future tense is indicated by the present tense on the surface.

Taking into account the actual tense of the relative clause instances, we calculated the following score:

$$V_{past} \begin{cases} R_{past} - AVG_{past} & \text{in case of present tense} \\ AVG_{past} - R_{past} & \text{in case of past tense} \end{cases} \quad (6)$$

For a head noun not implying the concept of time, in either tense of the main verb, the score is rather low, and a decision on inner/outer might not be affected by the score. For a head noun implying the concept of time, the absolute value of the score is rather large, and if the tense of the main verb is the same as the time concept, the score becomes negative; otherwise the score becomes positive.

8. Whether main verb has a sense of ‘exclusion’

The last feature is for identifying exceptional ‘outer’ relative clause. Consider the following two examples:

- (a) 日本を 除く アジア諸国
 Japan except Asian countries
 “Asian countries except Japan”
- (b) 怪我人を 除いた 乗客
 injured people except passenger
 “the passenger except injured people”

These examples are ‘outer’ relative clauses, and this RCC type is identified by the main verb which has sense of exclusion. There are a few verbs which indicate the RCC type by itself. Therefore, we defined a feature ‘excluding verb’. If the main verb contains a character ‘除’ (which has sense of exclusion), the feature is set to 1, otherwise, 0.

5 Machine Learning Based Classifier for RCC Type

We integrated the eight features in described the last section and used the machine learning approach to determine the RCC type. We used C5.0[13] as the machine learning algorithm.

C5.0 is a decision-tree based classification system that has been used in natural language processing, such as text classification, chunking, text summarization, and ellipsis resolution[10]. C5.0 takes a set of training instances with a feature vector and correct type as input, and induces a classifier which characterizes the given feature space.

Since we use only eight features, we think even the state of the art machine learning method like SVM would yield almost the same accuracy as decision-tree. Furthermore decision-tree are more easily interpreted by human than SVMs.

6 Evaluation

6.1 Experiment

Cooccurrence and other statistical information used in this work were calculated from the corpus of a collection of twenty-two years of newspaper articles. The corpus was parsed with KNP[8], which is a rule-based Japanese syntactic parser. The cooccurrence information we obtained was as follows: the number of $f_k(n, v)$ was about 60.8 million, and the number of $f_m(n, v)$ was about 12.4 million.

The data used in the evaluation was a set of RCC instances randomly extracted from the EDR corpus[4] which had syntactically analyzed. Then, a label, whether the relative clause is ‘inner’ or ‘outer’, was manually annotated. The statistics on the data are shown in Table 7. Evaluation with C5.0 was carried out by way of 5-fold cross validation.

Table 7. Statistics on evaluation data

Total	Inner	Outer
749	580	169

Table 8. Experimental results

	accuracy	Inner		Outer	
		precision	recall	precision	recall
Baseline	0.774	0.774	1.000	-	-
Cooccurrence information only	0.787	0.836	0.906	0.520	0.366
Case frame	0.830	0.868	0.921	0.657	0.521
Our approach	0.902	0.931	0.942	0.794	0.762

```

...excluding verb = 1: outer(exceptional type) (22/2)
:.excluding verb = 0:
  ..outerness <= 0.212: inner (444/6)
  outerness > 0.212:
    ...other modifier = 1: inner (84/17)
    other modifier = 0:
      ..cooccurrence("を" case) > -9.10: inner (28/4)
      cooccurrence("を" case) <= -9.10:
        ..percentage of "という" > 0.027: outer (105/14)
        percentage of "という" <= 0.027:
          ...percentage of modified <= 0.735: inner (25/2)
          percentage of modified > 0.735:
            ...cooccurrence("が" case) <= -13.1:outer (31/5)
            cooccurrence("が" case) > -13.1:inner (10/2)

```

Fig. 1. Generated decision tree

The baseline we used determines all instances as ‘inner’ relative clauses. We also compared our approach with the traditional method with case frames, and a method that uses only cooccurrence information (features 2 and 3 in section 4). An evaluation measure is an accuracy, which is defined as the number of correctly identified RCCs divided by the number of all RCCs. And for inner/outer relative clauses, precision and recall are calculated.

$$Precision = \frac{\# \text{number of correctly identified relative clauses}}{\# \text{number of inner/outer attempted by system}}$$

$$Recall = \frac{\# \text{number of correctly identified relative clauses}}{\# \text{number of inner/outer relative clauses}}$$

The results are shown in Table 8. The generated decision tree from all instances is shown in Figure 1. The last values on each line, for example ‘22/2’ and ‘444/6’, indicated ‘number of applied examples / number of misclassification’.

6.2 Discussion

Accuracy of our approach is higher than that of the traditional approach. Our approach works well especially for identifying ‘outer’ relative clause. Furthermore, using only cooccurrence information could not yield better performance for ‘outer’ relative clause. Therefore, we conclude that the features in our approach can effectively identify the ‘outer’ relative clause.

Figure 1 shows that the most contributive feature except ‘excluding verb’ is the degree of ‘outerness’. This feature can classify many instances with high accuracy (98.6%=438/444). If the degree of ‘outerness’ is smaller than certain threshold, RCC type is ‘inner’ with high probability.

The second contributing feature is the ‘other modifier’. If modifiers other than the relative clause exist, RCC type is ‘inner’. However, the accuracy of this feature is not so good compared with other features.

We unfortunately could not find the ‘tense information’ in our decision tree. We consider the reason to be that nouns which imply the concept of time are very few, and there might be no instances which contain them.

7 Conclusions

In this paper, we presented eight lexical and semantic features that characterized RCC, and we integrated them using machine learning approach to determine the RCC type.

Evaluation proved that our approach outperformed the traditional case frame-based method, and the features that we presented were effective in classifying types into ‘inner’ and ‘outer’ relative clauses.

After identification of ‘inner’ clauses, case identification will be necessary for semantic analysis. This will be considered in future work.

References

1. Baldwin, T., Tokunaga, T. and Tanaka, H.: The parameter-based analysis of Japanese relative clause constructions. In *IPSJ SIGNote on Natural Language* **134-8** (1999) 55-62
2. Baldwin, T.: Making Sense of Japanese Relative Clause Constructions. In *Proceedings of the Second Workshop on Text Meaning and Interpretation* (2004) 49-56.
3. Dagan, I., Lee, L. and Pereira, F.: Similarity-based models of word cooccurrence probabilities. *Machine Learning* **34** (1999) 65-81
4. EDR.: EDR electronic dictionary technical guide. Technical Report TR045, Japanese Electronic Dictionary Research Institute Ltd (1995)
5. Ikehara, S., Shirai, S., Yokoo, A. and Nakaiwa, H.: Toward an MT system without pre-editing effect of new methods in ALT-J/E . In *Proceedings of the Third Machine Translation Summit* (1991)
6. Ishizako, T., Kataoka, A., Masuyama, S., Yamamoto, K. and Nakagawa, S.: Reduction of overlapping expressions using dependency relations. *Natural Language Processing* **7**(4) (2000) 119-142. (in Japanese)
7. Kawahara, D. and Kurohashi, S.: Fertilization of case frame dictionary for robust Japanese case analysis. In *Proceedings of the 19th International Conference on Computational Linguistics* (2002) 425-431
8. Kurohashi, S. and Nagao, M.: Kn parser: Japanese dependency/case structure analyzer. In *Proceeding of the International Workshop on Sharable Natural Language Resources* (1994) 48-55
9. Lin, J.: Divergence measures based on the shannon entropy. *IEEE TRANSACTIONS ON INFORMATION THEORY*. **37**(1) (1991) 145-151
10. Manning, C. and Schutze, H.: *Foundations of Statistical Natural Language Processing*. MIT Press (1999)
11. Murata, M.: Extraction of negative examples based on positive examples automatic detection of mis-spelled Japanese expressions and relative clauses that do not have case relations with their heads . In *IPSJ SIGNote on Natural Language* **144-15** (2001) 105-112. (in Japanese)
12. Narita, H.: Parsing Japanese clauses modifying nominals. In *IPSJ SIGNote on Natural Language* **99-11** (1994) 79-86. (in Japanese)
13. Quinlan, J.: *C4.5: Programs for Machine Learning*. Morgan Kaufmann (1993)
14. Teramura, H.: *Rentai-shuushoku no shintakusu to imi*. No.1-4. *Nihongo Nihonbunka* **4-7** (1975-1978) (in Japanese)