Classification of Adjectival and Non-adjectival Nouns Based on their Semantic Behavior by Using a Self-Organizing Semantic Map

Kyoko Kanzaki Qing Ma Masaki Murata Hitoshi Isahara

Computational Linguistics Group
Communications Research Laboratory
{kanzaki, qma, murata, isahara}@crl.go.jp

Abstract

We treat nouns that behave adjectively, which we call adjectival nouns, extracted from large corpora. For example, in "financial world" and "world of finance," "financial" and "finance" are different parts of speech, but their semantic behaviors are similar to each other. We investigate how adjectival nouns are similar to adjectives and different from non-adjectival nouns by using self-organizing semantic maps. We create five kinds of semantic maps, i.e., semantic maps of abstract nouns organized via (1) adjectives, (2) adjectival nouns, (3) non-adjectival nouns and (4) adjectival and adjectival nouns and a semantic map of adjectives, adjectival nouns and non-adjectival nouns organized via collocated abstract nouns, and compare them with each other to find similarities and differences.

1 Introduction

In this paper, we propose a method for fundamental research to construct an organized lexicon, in which we classify words depending on not only their part of speech, but also their semantic categories. We applied both a neural network model and a linguistic method, that is syntactic information, to a large corpora and extracted necessary information. To extract semantic information of words such as synonyms and antonyms from corpora, previous research used syntactic structures (Hindle 1990, Hatzivassiloglou 1993 and Tokunaga 1995), response time to associate synonyms and antonyms in psychological experiments (Gross 1989), or extracting related words automatically from corpora (Grefensette 1994). Most lexical classification is based on parts of speech, as they have very important semantic information. For examples, typically, an adjective refers to an attribute, a verb refers to a motion or an event, and a noun refers to an object. However, in real data, a semantic function of a part of speech is not defined rigidly, as shown in the above examples. In spite of different parts of speech, they sometimes represent the same or very similar semantic functions. For examples, there are the following Japanese examples:

yuushuu_na *seiseki*
(excellent) (an academic record)
an excellent academic record

sugure_ta *seiseki*
(excel and suffix of "adnominal") (an academic record)
an excellent academic record

"*Yuushuu_na* (excellent)" is an adjective and "*sugure_ta* (excel)" is a verb, but they represent the same meaning and same semantic function, that is, an evaluation of an academic record. In English there are the following examples;

financial world
world of finance

In these examples, "financial" and "finance" are different part of speech, but represent same meaning and same semantic function, that is, one of domains.

On the other hand, there are examples in which only semantic function is the same, but the part of speech and meaning of the words are different. For examples,

kandai_na *kihuu* *no* *hito*
(gentle) (disposition) (of) (person)
a gentle person

shinshu *no* *kihuu* *no* *hito*
(initiative) (of) (disposition) (of) (person)
a person of initiative

In Japanese "*kandai_na* (gentle)" is an adjective and "*shinshu* (initiative)" is a noun. They have different parts of speech and meanings, but the same semantic function, that is, they represent characteristics of a person. In terms of a semantic function of representation of characteristics, both "*kandai_na* (gentle)" and "*shinshu* (initiative)" are classified in the same category. In this work we call this type of noun an "adjectival noun."

It is important for developing high quality natural language processing systems to establish an objective method to represent relationship between words not only by part of speech but also by semantic functions. However, it is very difficult to extract this type of linguistic phenomena from real data automatically. We used syntactic and semantic patterns in our previous work (Isahara and Kanzaki 1999) in order to extract these types of examples from large corpora semi-automatically. In this work, by using syntactic information, we are collecting adjectives and adjectival nouns in the "noun + NO (of + Noun)" structure that we supposed to have the same semantic functions. We examined how adjectives and adjectival nouns extracted from corpora are similar or different in the real data and how non-adjectival nouns unlike adjectival nouns are different from adjectives in order to verify the usefulness of self-organizing semantic maps for lexical semantics.

In Section 2, we explain our methodology, based on linguistic information. In Section 3, we describe a self-organizing semantic map. In Section 4, we describe the similarities between adjectives and adjectival nouns and the differences between adjectival nouns and non-adjectival nouns by comparing two different self-organizing semantic maps. In Section 5, we give our conclusion.

2 Methodology

Isahara and Kanzaki (1999) classified semantic relations between adjectives and their head nouns from the viewpoints of syntax, semantics and computational treatment. Among various types of semantic relations extracted in this research, there is a case in which the meanings of adnominal constituents are semantically similar to the features of their head nouns. Let us consider the Japanese phrases, "*kanashii kimochi* (sad feeling)" and "*yorokobi no kimochi* (feeling of delight)" as examples.

kanashii	*kimochi*
(sad)	(feeling)
{adjective}	{noun}

sad feeling

yorokobi	*no*	*kimochi*
(delight)	(of)	(feeling)
{noun}	{postpositional}	{noun}

feeling of delight

(The English translation of the "noun + *no*" examples should be read from right to left.)

One meaning of "*kimochi* (feeling)" represents the semantic element, [mental state]. In the above examples, the adjective, "*kanashii* (sad)," and "noun + *no*" structure, "*yorokobi no* (delight + *no*)," represent the concrete contents of their head noun "*kimochi* (feeling)," i.e. they are descriptors of the mental state: "*kimochi* (feeling)." The head noun, "*kimochi* (feeling)," is a cognate object for "*kanashii* (sad)" and "*yorokobi no* (delight + *no*)." Therefore, even though "*kanashii* (sad)" and "*yorokobi no* (delight + *no*)" belong to different parts of speech (adjective and noun phrase), they must be classified as the same semantic category, since both carry the same type of meaning.

As for data, necessary expressions are extracted from large corpora: 10 year's worth of Japanese newspapers — the *Mainichi Shinbun* from 1991 to 2000, 100 novels — *Shincho-bunko*, and 100 kinds of essays. We extracted 134 abstract nouns used as this kind of head noun semi-automatically by using syntactic patterns that Isahara and Kanzaki(1999) and Kanzaki et al. (2000) used in their paper. The total number of adnominal constituents appearing with these head nouns in the corpora was 47,248, and the number of different adnominal constituents was 28,063. We got the list of pairs of a head (abstract) noun and its adnominal constituents (Table 1). These adnominal constituents are classified into three types, i.e. adjectives, adjectival nouns and non-adjectival nouns.

Table 1: Example of gathered data

Noun	Adnominal constituents
kimochi (feeling)	*shiawasena* (happy), *hokorashii* (proud), *kanashii* (sad), …
joutai (status)	*aimaina* (vague), *ansei no* (repose + *no*), …
kanten (viewpoint)	*gakumontekina* (academic), *anzensei no* (safety + *no*), …
…	…

We classified these head nouns according to the similarities of sets of their adnominal constituents by using a self-organizing system in a neural network model. This means that we co-classified both head nouns, i.e. abstract nouns, and adnominal constituents at the same time.

3 Self-Organizing Semantic Map

In this section, we explain self-organizing semantic maps by a neural network model. For the analysis of the similarities between adjectives and adjectival nouns, we make some semantic maps based on these adnominal constituents. We use a self-organizing semantic map to classify words because it distributes words onto a two-dimensional plane and is therefore a visible and continuous representation. This feature is very feasible to classify word meanings, because they cannot be always classified into an explicit category as hierarchical clustering does. As for the clustering ability of self-organizing semantic map compared with the multivariate statistical analysis and hierarchical clustering method, it is almost the same as the hierarchical clustering method and superior to multivariate statistical analysis (Ma 2001).

The semantic map we construct in this paper is one on which nouns, with their adnominal constituents as attributes, are mapped in a semantic order; i.e. nouns with similar meanings are mapped on (i.e. best-matched by) nodes that are topographically close to each other, and words with meanings that are far apart are mapped on nodes that are topographically far apart.

3.1 Learning Data

As we mentioned above, we used the list in Table 1 as learning data. Table 1 shows some example data that was gathered manually and in which the adnominal constituent is a descriptor of its head noun, i.e. a kind of cognate noun.

3.2 Encoding

The semantic map of nouns is constructed by first defining each noun as the set of its adnominal constituents. From Table 1, for example, we can define "*kimochi* (feeling)" as the set of its adnominal constituents, i.e. "*kimochi*" = {"*shiawasena* (happy)," "*hokorashii* (proud)," "*kanashii* (sad)," "*kinodokuna* (unfortunate),"…}. Suppose there is a set of nouns w_i (i = 1, … , ω) that we are planning to use for self-organizing. Any noun w_i can be defined by a set of its adnominal constituents as

$$w_i = \{ a_1^{(i)}, a_2^{(i)}, \ldots, a_{\alpha_i}^{(i)} \} \qquad \text{-------(1)}$$

where $a_j^{(i)}$ is the jth adnominal constituent of w_i and α_i is the number of adnominal constituents of w_i. One method of encoding nouns so that they can be treated by SOM is to use random coding, which is a common method used for constructing SOMs (see details in Kohonen (1997)). By several preceding computer experiments, however, we found that this method is not suitable for our task. We therefore used a new method as described below.

Suppose we have a correlative matrix (Table 2) where d_{ij} is some metric of correlation (or distance) between nouns w_i and w_j. We can encode noun w_i from the correlative matrix as

$$V(w_i) = [d_{i1}, d_{i2}, \ldots, d_{i\omega}]^T. \qquad \text{---------(2)}$$

The $V(w_i) \in \Re^{\omega}$ is the input to the SOM, i.e. $x = V(w_i)$ and $n = \omega$.

Table 2: Correlative matrix of nouns

	w_1	$w_2 \ldots$	w_ω
w_1	d_{11}	$d_{12} \ldots$	$d_{1\omega}$
w_2	d_{21}	$d_{22} \ldots$	$d_{2\omega}$
⋮	⋮	⋱	
w_ω	$d_{\omega 1}$	$d_{\omega 2} \ldots$	$d_{w\omega}$

In this paper, d_{ij} is measured by

$$d_{ij} = \begin{cases} \dfrac{(\alpha_i - c_{ij}) + (\alpha_j - c_{ij})}{\alpha_i + \alpha_j - c_{ij}} & \text{If } i \neq j \\ 0, & \text{otherwise} \end{cases} \qquad \text{-------(3)}$$

where α_i and α_j are respectively the numbers of the adnominal constituents of w_i and w_j, and c_{ij} is the total number of common adnominal constituents of both w_i and w_j. The term d_{ij} is therefore a normalized distance between w_i and w_j in the context of the number of adnominal constituents they have in common; the smaller d_{ij} is, the closer w_i and w_j are in terms of their adnominal constituents.

4 Experimental Result

4.1 Comparisons of Word Distribution on Semantic Map via Adjectives with ones via Adjectival Nouns and via Non-adjectival Nouns.

In Section 4, we examine adjectival nouns extracted from corpora, whose behaviors are similar to adjectives. In order to verify the data extracted from corpora manually by using syntactic method we prepare for four kinds of self-organizing semantic maps. One is a semantic map of head nouns co-occurring with adjectives the second is a semantic map of head nouns co-occurring with adjectival nouns that we extracted from corpora, the third is a semantic map of head nouns co-occurring with non-adjectival nouns and the final one is a semantic map of head nouns co-occurring with both adjectives and adjectival nouns. As we mentioned in section 2, head nouns distributed on four maps are abstract nouns that represent the concrete content of adnominals, e.g., “feeling” co-occurring with “happy” and so on.

We compare a semantic map of head nouns via co-occurring adjectives (Figure 1) with three other maps, that is, a semantic map via adjectival nouns and a semantic map via non-adjectival nouns and a semantic map via adjectives + adjectival nouns. And then we mark the points of words that are similarly distributed between the semantic map via adjectives and one of other maps (Figure2, 3, 4).

Input data for neural network model was a list that we mentioned in section 2. In the ordering phase, the number of learning steps was 10,000, and in the adjustment phase it was 100,000. After learning the input data, a two-dimensional array, in which a hexagonal topology type of neighborhood, the area that the winner node influences in the learning stage was used. A self-organizing semantic map of head nouns via adjectives is shown in Figure 1. We translate some Japanese words on the map into English for the reader’s convenience.

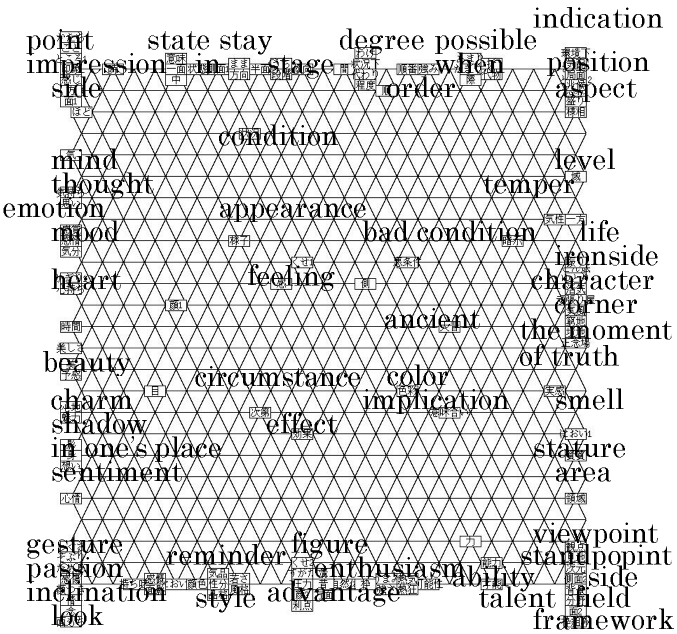

Figure 1. Self-organizing semantic map of head nouns via adjectives

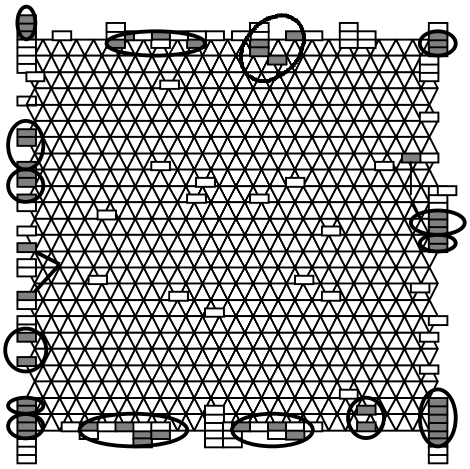

Figure 2. A semantic map via adjectives marked via comparison with classification by adjectival nouns.

For examples, “viewpoint,” “standpoint,” “side” on the right hand in Figure 1 are co-occurring with “medical,” “musical,” “economical,” “political” and so on, and “mind,” “thought,” “mood” are co-occurring with “delightful,” “sad,” “happy,” “proud” and so on. Some sets of head nouns are not classified enough because the number of the co-occurring adjectives is not enough, or we could not extract enough of the collocation that we treated.

After we made a semantic map of head nouns via adjectival nouns we compared it with a semantic map of head nouns via adjectives (Comparison1). We marked the words in Figure 1 located similarly between two semantic maps (See Figure 2). We defined “the common sets of words” as words located similarly between two semantic maps, i.e., marked words, and they are located within three neighborhoods on the semantic map (See also Figure 2).

And we also examined the data of non-adjectival nouns as same as the above experience and we marked the common sets of words between two maps, that is, a semantic map via non-adjectival nouns and a semantic map via adjectives (Comparison 2).

Each square on Figure 2 and 3 refers to a word in Figure 1. The black marked squares indicate a common words appearing on both a map via adjectives and a map via another data and a circle surrounding squares is common sets of words on both maps.

In Figure 2 we marked 51 words among 134 abstract nouns (38% of all the abstract nouns) on a semantic map via adjectives (Figure 1), which are common in a classification of words on two self-organizing semantic map, i.e., two semantic maps via adjectives and via adjectival nouns and these 51 words can be classified into 16 common sets of words.

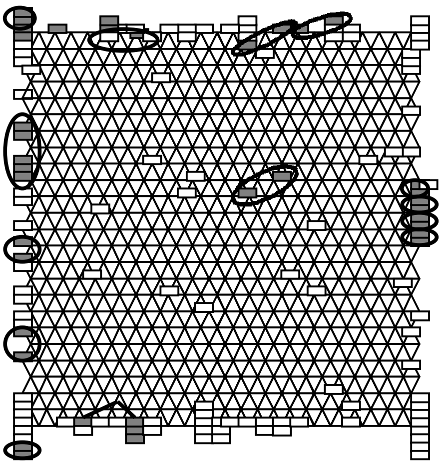

Figure 3. A semantic map via adjectives marked via comparison with classification by non-adjectival nouns.

Then, we compared the semantic maps organized via non-adjectival nouns with semantic map via adjectives to find how different these maps are. Thirty-five marked words from 134 abstract nouns (26%) and 14 common sets of these words, which are common between two maps, i.e., semantic maps via adjective and via non-adjectival nouns, are distributed on semantic maps via adjectives (Figure 3).

There are 12% more common words and 3 more common sets in Comparison 1 than in Comparison 2. However, there is a question of why the map organized via non-adjectival nouns still has sets of words common to the map organized via adjective. Are there any similarities of behaviors between adjectives and non-adjectival nouns? We investigated the common co-occurring head nouns in Comparison 2 precisely, and found two facts that caused the existence of these common sets of words in Comparison 2.

One is that some co-occurring words that we classified as non-adjectival nouns are nouns that we must classify as adjectival nouns. Another is that some non-adjectival nouns refer to people and they are possessors of the modified abstract nouns. For examples, “emotion,” “mood” and “thought” are common sets in both maps. Co-occurring adjectives are “delight,” “sad” and “happy,” however, co-occurring non-adjectival nouns are “*watashi-no* (my),” “*haha-no* (mother’s),” “sensei-no (teacher’s),” and so on. From this fact, we can conclude that the existence of common classifications of head nouns between these two semantic maps does not always mean semantic similarity between adjectives and non-adjectival nouns.

From these observations we made a semantic map of head nouns by using both adjectives and adjectival nouns. If the adjectival nouns work similarly to the adjectives, using both adjectives and adjectival nouns will not influence the distribution and classification of words on the semantic map via adjectives. On the other hand, if the data of semantic phenomena between adjectives + adjectival nouns and adjectives only are completely different, the distribution and classification of words on the semantic map via adjectives will be influenced by the addition of adjectival nouns. We mark the point of the common words between them on the semantic map via adjectives (Figure 4).

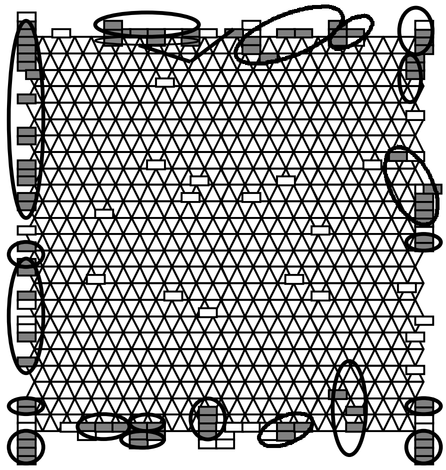

Figure 4. A semantic map via adjectives marked by the common words between classification by adjectives and adjectival nouns and classification by only adjectives.

Eighty-three words among 134 words on this map are classified similarly to the words on the map organized via adjectives, and there are 21 similar sets of words. This result shows that the distribution of the abstract nouns on the semantic map is not affected by the addition of adjectival nouns. Therefore, the semantic roles of adjectival nouns for abstract nouns are similar to those of adjectives.

4.2 A Semantic Map Distributed by Adjectives, Adjectival Nouns and Non-adjectival Nouns Organized via Head Nouns.

In this section we made the semantic map of adjectives, adjectival nouns and non-adjectival nouns organized via collocation with abstract nouns to see the semantic distances between them. As for marks on the map, ◆, ● and ◎ indicated, in turn, adjective, adjectival nouns, and non-adjectival nouns.

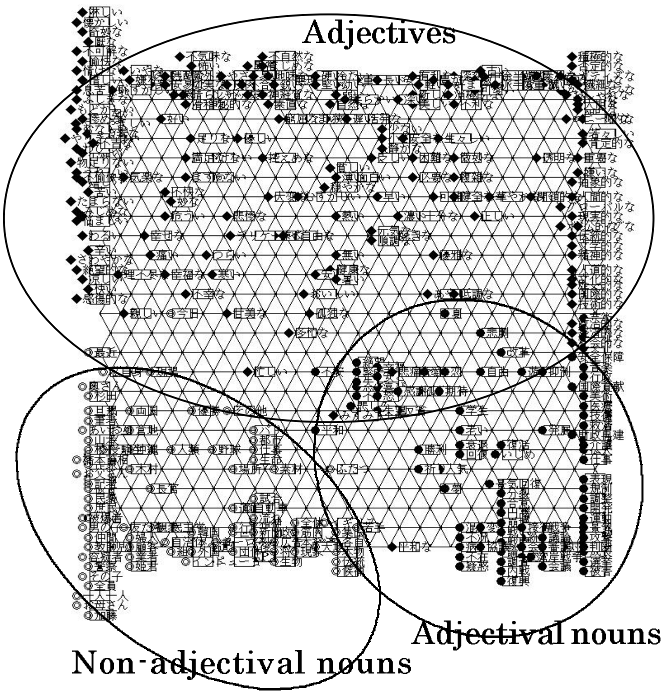

Figure 5. Semantic map of adjectives, adjectival nouns and non-adjectival nouns organized via collocation with abstract nouns.

We distributed three kinds of words, that is, adjective, adjectival nouns and non-adjectival nouns on the semantic map based on their head nouns, that is, abstract nouns. For example, “happy” has “feeling,” “mood,” “state,” and so on as co-occurring head nouns. When we made this map, we utilized words (adjectives, adjectival nouns and non-adjectival nouns) that collocate with 10 to 20 abstract nouns, so that the input data for constructing semantic map is fair from the viewpoint of number of co-occurring words. We selected from them 100 adjectives, 100 adjectival nouns, and 200 non-adjectival nouns at random. This semantic map is shown in Figure 5.

The semantic map shown in Figure 5 shows that there are three classes on the map. The upper half part of this semantic map indicates the adjective area, the bottom right half of this map is the adjectival noun area and the bottom left half of this map is the non-adjectival noun area. Semantic roles of adjectives are isolated from those of nouns, and semantic roles of nouns are divided into two areas, i.e. adjectival and non-adjectival. The self-organizing mechanism could separate the semantic roles of adjectival nouns from those of non-adjectival nouns.

5 Conclusion

We extracted adnominal constituents from corpora and created several self-organizing semantic maps by using them.

First, we compared the semantic maps organized via adjectives and via adjectival nouns. The common sets of head nouns were 16 sets and common head nouns were 51 words in 134 head nouns, that is, 38% of the head nouns were classified similarly.

Second, we compared the semantic maps organized via adjectives and via non-adjectival nouns. The common sets of head nouns were 14 sets, and the common head nouns were 35 words in 134 head nouns, that is, 26% of the head nouns were the same classifications.

Some sets of abstract nouns, head noun co-occurring with adjectives, are common with sets of abstract nouns co-occurring with non-adjectival nouns. However, based on the precise investigation, we could find that the semantic function of adjectives and non-adjectival nouns were different.

Finally, we created a semantic map of abstract nouns by both adjectives and adjectival nouns. This is because we wanted to see how word distribution on the map changed when we added adjectival nouns to the data for self-organization. The common sets of head nouns were 21 sets and the common head nouns that did not change were 83 words in 134 abstract nouns, that is, 62% of head nouns were not affected by the addition of adjectival nouns. This means that adjectival nouns are similar to adjectives in their semantic behavior for abstract nouns.

Then, we showed the semantic map of adjectives, adjectival nouns and non-adjectival nouns organized via co-occurring abstract nouns. As these three kinds of adnominals were isolated on this map, we could find that the adjectival nouns had specific semantic roles that are different from those of non-adjectival nouns.

From the above evidence, we considered that we could extract the adjectival nouns similar to adjectives, rather than non-adjectival nouns.

In future work, we need addition and modification of input data and would like to use the accurate distribution of words by using some kind of information such as frequencies[1]. And then we will construct a semantic map of words from Japanese large corpora and link words according to semantic behavior while we verify our data extracted from corpora by using a neural network model.

References

Hindle, D. (1990), *Noun Classification From Predicate-Argument Structures*, In the Proceedings of the 28th Annual Meeting of the Association for Computational Linguistics, *pp. 268-275.*

Hatzivassiloglou, V. and McKeown, R.K. (1993), *Towards the Automatic Identification of Adjectival Scales: Clustering Adjectives According to Meaning,* In the Proceedings of the 31st Annual Meeting of the Association for Computational Linguistics, *pp. 172-182.*

Tokunaga, T. Iwayama, M. and Tanaka, H. (1995), *Automatic Thesaurus Construction Based On Grammatical Relations*, In the Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI), *pp. 1308-1313.*

Gross, D. (1989), *The Organization of Adjectival Meanings,* In Journal of Memory and Language, vol. 28, *pp. 92-106*

Grefenstette, G. (1994), Explorations in Automatic Thesaurus Discovery, Kluwer Academic Publishers.

Isahara, H. and Kanzaki, K (1999) *Lexical Semantics to Disambiguate Polysemous Phenomena of Japanese Adnominal Constituents*, ACL ’99.

Kanzaki, K., Ma., Q. and Isahara, H. (2000), *Similarities and Differences among Semantic Behaviors of Japanese Adnominal Constituents,* Workshop on the Syntactic an Semantic Complexity in Natural Language Processing Systems, ANLP and NAACL.

Ma, Q. Kanzaki, K., Murata, M., Uchimoto, K. and Isahara, H. (2001) *Self-Organizing Semantic Map of Japanese Nouns*, Information Processing Society of Japan, Vol. 42, No. 10.

[1] We can use many Japanese newspapers as corpora, however we cannot use many kinds of texts, for examples, novels, essays and so on. Though we should try to use frequencies, if our corpora are not well balanced, we don’t know how much confidence to place on the result of frequencies. This is why we didn’t use the frequencies as the first step of our research.