A Disambiguation Method for Japanese Compound Verbs

Kiyoko Uchiyama and Shun Ishizaki
Graduate School of Media and Governance
Keio University
5322 Endo, Fujisawa-shi, Kanagawa, 252-8520, JAPAN
{kiyoko, ishizaki}@sfc.keio.ac.jp

Abstract

The purpose of this study is to construct a semantic analysis method for disambiguating Japanese compound verbs. Japanese speakers produce a rich variety of compound verbs, making it difficult to process them by computer. We construct a method employing 110 disambiguation rules based on the semantic features of the first verb of a compound and syntactic patterns consisting of co-occurrence between verbs and nouns. The disambiguation rules are evaluated by applying them to compound verbs in the dictionary. The obtained accuracy is 87.19% for our rules. This result shows the advantage of our method.

1 Introduction

The treatment of multiword expressions (MWEs) has attracted much interest as an important issue (Sag et al. 2002). Japanese has diverse types of MWEs and there are difficulties in processing them (Baldwin and Bond 2002). In those studies, MWEs are defined as a "word with spaces" in English and "idiosyncratic interpretations that cross word boundaries" in Japanese. Verb particles as one type of MWE in English have been studied (Villavicencio and Copestake 2002). We predict that verb particles in English and compound verbs in Japanese have commonalities in terms of ambiguity and semantic constraints. For example, the English particle "up" has both an aspectual ("finish up writing" cf. "*kaki-owaru*") and a spatial meaning ("go up the stairs" cf. "*kake-agaru*"), which is equivalent to the second verb in Japanese compound verbs. We investigate Japanese compound verbs (JCVs) and extract semantic constraints for the purpose of applying them to machine translation.

JCVs consist of two verbs, the first verb (V1) and the second verb (V2). V1 always appears in the continuative form. In this paper, we discuss only "Verb-Verb" JCVs, which are a composition of two native Japanese verbs, for example *oshi-ageru* "push up", *tabe-sugiru* "eat too much" and so on. JCVs of V1-V2 form are frequently used for expressing complex motion, elaborated phenomena and emotional state. However, JCVs have high productivity, a great number of ambiguities and semantic constraints between each constituent. Examples of correlation between semantic constraint and ambiguities of JCVs are given in (1).

(1) a. *nage-ageru* "throw up", *keri-ageru* "kick up", *mochi-ageru* "lift up", *oshi-ageru* "push up"
b. *yude-ageru* "finish boiling", *mushi-ageru* "finish steaming", *yaki-ageru* "finish baking"

Ageru "lift" has multiple meanings when it appears in the V2 position of a JCV. A directional compound verb in (1a) is formed by compounding *ageru* "lift" as V2 with a verb of motion like *nageru* "throw" and *keru* "kick" as V1. Conversely, an aspectual compound verb in (1b) is formed by combining *ageru* "lift" as V2 with a verb of cooking process like *yuderu* "boil" and *musu* "steam" as V1. There are a small number of such ambiguous verbs which appear as V2. Ambiguous JCVs are generated by compounding various instances of V1 with an ambiguous V2, which makes it difficult to process on computer.

The analysis of JCVs has been discussed in the field of linguistics and natural language processing.

In linguistics, JCVs have been studied mainly in terms of syntax (Kageyama 1999) and constraints on semantic structures (Matsumoto 1996, 1998). Himeno (2000) made a semantic analysis concerning the types of V2s which have multiple meanings. She classified JCVs by the meaning of their V2. However, she confounded the meaning of V1 with that of V2 in her classification. In order to clarify the semantic constraints between V1 and V2, we need to analyze each constituent individually.

In natural language processing, Shirai (1998) proposed a method of building valency patterns for JCVs by compiling a Japanese and English corpus. This approach, in which whole compound verbs are registered in an electronic dictionary, can improve the translation rate of the system. However, it is inefficient to register all the compound verbs in advance. For that reason, it is desirable to develop a framework for understanding JCVs by processing each constituent.

Based on this background, we propose a method employing rules which utilize semantic features and syntactic information to clarify the semantic constraints for disambiguation of JCVs.

In this paper, we take two steps in order to construct a disambiguation method. The first step is to identify the meaning of V1 using by rules which an MT system or other lexical database should have already. The second step is to classify JCVs into semantic clusters and extract commonalities of semantic features on V1 (semantic information) and verb complements (syntactic information). We build rules using this obtained semantic and syntactic information. This is the major innovation of this study. The proposed method based on disambiguation rules has the advantage of being able to analyze new compound verbs not in the dictionary. Since the semantic restrictions of JCVs are similar to those of phrasal verbs in English (Villavicencio and Copestake 2002), there might be a possibility of applying our method to machine translation.

The rest of the paper is structured as follows. Section 2 describes the definition, ambiguities and semantic relations of JCVs. Section 3 shows analysis results. The semantic analysis method is explained in Section 4. The evaluation of this method is discussed in Section 5. The conclusion of our study and implication for future work are stated at the end.

2 Ambiguities of JCVs

2.1 Types of Ambiguities

Kageyama (1993) has proposed that JCVs can be analyzed by the argument structure of each constituent and divided into two types: syntactic compounds and lexical compounds. Lexical compounds have semantic constraints and are limited to lexically specified combinations, whereas syntactic compounds are basically compositional and have no lexical idiosyncrasies .

We do not differentiate these two types in advance, because our method may be also useful for identifying them. There are two types of ambiguity in JCVs: ambiguities within lexical compounds and ambiguities between lexical compounds and syntactic compounds.

Lexical compounds containing an ambiguous V2 (as in example (1)) are examined in this study. Semantic constraints govern the pairs of verbs which may be compounded. The semantic features of V1 play a key role in identifying the meaning of V2. We focus on extracting commonalities of semantic features from V1 in order to disambiguate V2.

On the other hand, some JCVs are ambiguous, because they may be either syntactic or lexical compounds depending on context. Syntactic information is important in disambiguating this type of JCV. Example (2) indicates that JCVs with the same morphology change their meanings depending on specific context.

(2) a. *Basu wa basutei o hashiri-sugita* .
"The bus ran past the bus stop."
b. *Kare wa shiai no tame ni hashiri-sugita.*
"He ran too much because of the game."

V2 *sugiru* in the lexical compound in (2a) means path of motion ("go past"), but in the syntactic compound in (2b) it denotes excessiveness ("too much"). As sugiru is most commonly used as a compositional V2 ("too much"), it is difficult to identify the difference between (2a) and (2b). Since sentence (2a) includes a word indicating the place

like *basutei* "bus stop", we can distinguish the difference between the lexical compound (2a) and the syntactic compound (2b) by co-occurring words. We identify the meaning of such JCVs using syntactic information gained from co-occurrence and verb complements.

2.2 Ambiguities of V2

We classified ambiguities of V2 into three semantic clusters: **aspectual**, **spatial** and **adverbial** (Niimi 1987). An ambiguous V2 is defined as a word with multiple meanings which overlap several semantic clusters. This framework makes it easier to distinguish the difference in meaning for an ambiguous V2.

We listed all the ambiguous V2, i.e. the following 20 words, based on the previous study (Himeno 2001).

> *agaru* "go up", *ageru* "lift", *otosu* "drop", *kakeru* "hang", *kakaru* "hang onto", *kaeru* "go back", *kaesu* "send back", *iru* "enter", *komu* "insert", *sugiru* "go past", *tatsu* "stand", *tateru* "make stand", *tsuku* "be attached", *tsukeru* "attach", *dasu* "put out", *kiru* "cut", *toosu* "pierce", *nuku* "pull out", *tobasu* "scatter", *wataru* "go across"

For the first step of analysis, we extracted 10 ambiguous words at random, *agaru* "go up", *ageru* "lift", *otosu* "drop", *kakeru* "hang", *kakaru* "hang onto", *kaeru* "go back", *kaesu* "send back", *iru* "enter", *komu* "insert" and *sugiru* "go past". Table 1 shows examples of ambiguities of V2 as JCVs.

V2	Aspectual	Spatial	Adverbial
agaru "go up"	*yude-agaru* "finish boiling"	*tobi-agaru* "jump up"	*furue-agaru* "be terrified"
kakaru "hang onto"	*ochi-kakaru* "be dropping"	*kiri-kakaru* "slash at"	
otosu "drop"		*kiri-otosu* "cut off"	*ii-otosu* "forget to say"
kaeru "go back"		*uchi-kaesu* "hit back"	*waki-kaeru* "be highly excitted"
sugiru "go past"		*toori-sugiru* "go past"	*tabe-sugiru* "eat too much"
komu "insert"		*hairi-komu* "go into"	*fuke-komu* "become old"

Table 1. Types of ambiguities of V2

We are not concerned here with disambiguation of the meanings within a single cluster. For example, *tabe-kakeru* "already begin to eat" and *hashiri-kakeru* "be about to run" are both classified as members of the **aspectual cluster**. JCVs of the **adverbial cluster** which include *naosu* "fix" and *au* "fit" seem to be similar cases. Such differences are not analyzed in this study.

2.3 Criteria for Classification of Semantic Cluster

In order to classify JCVs into semantic clusters, we need to establish certain criteria. We define the syntactic roles and semantic relations between each constituent as criteria for classification into a semantic cluster. Each constituent of a JCV has syntactic roles such as dependency to a noun phrase and suffix usage. We refer to any JCV component verb that requires a complement as a main verb, and any suffixing component as a subsidiary verb.

We also need to examine how the two verbs are related each other within a JCV. The semantic relation classes are assigned to the JCV constituents respectively based on Tagashira's (1986) study. As the paraphrasing of V2 facilitates understanding of these semantic relations, we investigate them by paraphrasing.

The ambiguous JCVs were classified into three types based on the syntactic roles of V2: **complementation**, **modification** and **directional motion**. In **complementation**, V2 plays a complementary role to V1, and can be paraphrased using other aspectual words such as *hajimeru* "start" and *owaru* "finish". In **modification**, the V2 modifies, so we can paraphrase V2 with an adverb. **Directional motion** consists of two main verbs. The V1 expresses the manner of motion, and V2 the direction.

We describe the semantic relations of JCVs as "SEM", the syntactic roles of V1 as "SYN" and the paraphrase as "PAR", with examples as follows. The symbol '/' means "or". The criteria for classification into each semantic cluster are given in (3).

(3) Criteria for classification

a. **Aspectual cluster**

SEM: V1:motion/activity, V2:aspect

SYN: **Complementation**

(V1:main verb, V2:subsidiary verb)

PAR: 'V1 suru koto ga/o owaru/oeru/ hajimaru/hajimeru'

arai-ageru= arau koto o oeru "finish washing"

nomi-dasu= nomu koto o hajimeru "start to drink"

b. **Spatial cluster**

SEM: V1:manner, V2:motion

SYN: **Directional motion**

(V1:main verb, V2:main verb)

PAR: 'V1 shite V2 suru'

tori-dasu = totte dasu "take out"

c. **Adverbial cluster**

SEM: V1:motion/emotional state/activity

V2:intensity

SYN: **Modification/Complementation**

(V1:main verb, V2:subsidiary verb)

PAR: 'Hijouni (other adverb) V1 suru'

'V1 suru koto ga V2 suru'

tabe-sugiru= taberu koto ga sugiru "eat too much"

yomi-kaesu= mouichido yomu "read again"

3 Analysis of JCVs

In order to build disambiguation rules which are applicable to novel JCVs, we need to examine and analyze frequency, types of semantic features and co-occurring words of JCVs not in the dictionary.

3.1 Extraction of JCVs

We used data from the Mainichi Shinbun (Mainichi 1993) in order to examine JCVs not in the printed dictionary (Kindaichi 1999). The newspaper articles were tagged by the morphological analysis system Chasen (Matsumoto et al. 2000). All occurrences of "Verb-Verb" JCVs and non-compounded single verbs were extracted from the tagged articles. Table 2 shows the number of tokens extracted and the number of tokens after duplicates had been removed.

	Tokens	Types
"Verb"	1,437,123	6,810
"Verb-Verb"	18,689	5,364

Table 2. Tokens and types of JCVs

The "Verb-Verb" form accounts for only 1.36% of the total tokens, however, this type accounts for 44.06% of types. In addition, 3525 words (accounting for 65.71% of all "Verb-Verb" tokens) are not registered in the dictionary. The result shows a rich variety of JCVs and difficulty of processing JCVs using a static dictionary. 829 types of ambiguous JCVs, using of the 10 ambiguous V2s, mentioned in 2.2, were found in the 3525 JCVs not in the dictionary.

3.2 Semantic Features for Disambiguation of JCVs

The semantic features are necessary for representing appropriate meanings of V1. We used Ruigo Shin Jiten (Oono and Hamanishi 1989) to label the semantic feature. The framework of the semantic feature in Ruigo Shin Jiten provides enough accuracy to distinguish the meaning of V1. For instance, Ruigo Shin Jiten defines the semantic feature of *musu* "steam" as *suiji* "kitchen work" and that of *nageru* "throw" as *dageki* "throw and hit". These features can identify the different semantic clusters of *mushi-ageru* "finish steaming" in the **aspectual cluster** and *nage-ageru* "throw into the air" in the **spatial cluster**. Ruigo Shin Jiten is organized in three levels and constitutes 1000 categories. The labels from the second level, which include 60 categories for verbs, are used in assigning a semantic feature to V1. If it is difficult to identify the meaning of V1 using the label from the second level, the label from the third label is applied as the semantic feature.

4 Construction of Disambiguation Method

4.1 Information for Disambiguation Rules

We analyzed semantic features of all 328 V1s among 829 target words. We examined ambiguities of V1 based on the co-occurring nouns in a sentence using the syntactic information in IPAL verb dictionary (IPA 1987).

In order to construct disambiguation rules, we took two steps. The first step was to disambiguate the meaning of V1. The second step was to clarify semantic and syntactic information for use in the disambiguation rules.

As for the first step, we used the IPAL verb dictionary. The IPAL verb dictionary defines the meaning of verbs using valency patterns and assigns

a semantic feature from Ruigo Shin Jiten to each entry.

Initially, the co-occurring nouns and verb complements of JCVs were extracted from a sentence. For the purpose of examining the correlation between co-occurring nouns and V1, we investigate the valency patterns of V1 in the IPAL dictionary. For example, in the case of a sentence like *kare wa udon o uchi-ageta* "He completed making buckwheat noodle", we try to find the sub-entry of *utsu* "hit" having a complement such as '*kare wa*' and '*udon o*' in IPAL dictionary. When we can identify the sub-entry of *utsu* "hit" which fulfills this condition, a semantic label like *seisan* "production" is selected as the semantic feature for *utsu* "hit".

The second step is to classify JCVs into semantic clusters based on the criteria as defined in 2.3, and to extract commonalities of semantic features on V1 within the same semantic cluster. For example, JCVs such as *yude-ageru* "finish boiling", *mushi-ageru* "finish steaming" and *yaki-ageru* "finish baking" classified into the **aspectual cluster**, have a common semantic feature: *suiji* "house keeping".

Verb complements which are not related to V1 and their semantic features are used as syntactic information in disambiguating the meaning of V2. For instance, *unazuku* "nod" has a single meaning of "agreement". In combining *unazuku* "nod" with *kakeru* "hang" as V2, *unazuki-kakeru* causes two ambiguities. The first meaning is the aspectual meaning such as *kare wa sono kotoba ni unazuki-kaketa* "he was about to nod at what was being said." The second meaning is a spatial one such as *kare ni unazuki-kaketa* "I nodded at him". In this case, we need semantic features and syntactic information including verb complements.

4.2 Disambiguation Rules

In order to construct of disambiguation rules, the JCVs were classified into two groups using the results of the analysis in section 4.1.

The rules of the first group are based on the semantic features of V1. For example, *utsu* "hit" has two meanings, "hit" and "make". The semantic feature of *utsu* "hit" in the first meaning is labeled as *dageki* "hit and throw" in a specific context such as *kare wa bouru o uchi-ageta*: "he hit the ball up", and classified in **spatial cluster**. The second meaning is assigned *suiji* "cooking" as a semantic feature in a sentence such as *kare wa udon o uchi-ageta*: "he finished making buckwheat noodles", and categorized in **aspectual cluster**.

We built compounding rules for disambiguation utilizing the semantic features of Ruigo Shin Jiten. The rules are composed of the semantic features of V1 and verb of V2 and the corresponding semantic cluster.

Examples of these disambiguation rules are shown as follows.

Rules based on semantic information

Rule 1: IF V1 is-a cooking and V2 is '*ageru*'
THEN class is **aspectual cluster**
Example: *yude-ageru* "boil-raise"
=> **aspectual cluster**
Paraphrase: *yuderu-koto-o oeru*
"finish boiling"

Rule 2: IF V1 is-a operation and V2 is '*ageru*'
THEN class is **spatial cluster**
Example: *uchi-ageru* "hit-raise"
=> **spatial cluster**
Paraphrase: *utte-ageru* "hit upwards"

Rule 3: IF V1 is-a emotion and V2 is '*agaru*'
THEN class is **adverbial cluster**
Example: *furue-agaru* "tremble-go up"
=> **adverbial cluster**
Paraphrase: *hijouni furueru*
"tremble violently"

Rule based on syntactic information

Rule 4: IF V1 is-a action and
N1 (V1's subject) is-a human
and N2 (V2's dative) is-a human
and V2 is '*kakeru*'
THEN class is **aspectual cluster**
Example:
kare wa kanojo ni unazuki-kaketa
"he TOP she Goal nod-hang"
=> **aspectual cluster**
Paraphrase:

kare wa kanojo ni mukatte unazuita
"He nodded at her"

We extracted 829 JCVs from the newspaper articles as target words for the analysis. As a result of analyzing the 328 types of V1, 143 rules of semantic information and 35 rules of syntactic information were constructed for the disambiguation of JCVs.

4.3 Expanding of Disambiguation Rules

To expand the rules comprehensively in addition to the results in 4.1, we prepared a matrix based on our rules. Table 3 illustrates a part of a matrix we used.

Semantic features	*agaru* "go up"	*ageru* "lift"	*kaesu* "send back"	*kaeru* "go back"
sousa "operation"	+	+	+	-
seishin "mental state"	+	-	+	+
kenbun "communication"	-	-	+	-
te no dousa "motion by hand"	+	+	+	-
ongaku "music"	-	+	-	-
seisan "production"	+	+	-	-
hyojo "expression"	-	+	+	+

Table 3. A part of a matrix of disambiguation rules

The lists of V2 are shown in the row headings and the semantic features of V1 are described in the column. We verified the ability or inability of a V1 having semantic feature shown in the column to combine with the V2 in the row, and marked the ability with "+" and the inability with "-". The rules were compiled from the matrix and reconstructed. As the reconstruction reduced the number of rules, 110 disambiguation rules were obtained.

4.4 Disambiguation Method

We propose a semantic analysis method for JCVs based on disambiguation rules. The following 6 steps comprise our method.

(1) Input a sentence which includes JCVs
(2) Tag each word in the input sentence using a morphological analysis system called Chasen
(3) Extract the JCVs and their syntactic information from the sentence
(4) Assign a semantic feature to V1 using co-occurring words, referring to IPAL dictionary
(5) Compare the semantic feature of V1 and syntactic information with the disambiguation rules
(6) Output the semantic cluster obtained by application of the matching rule

Through this procedure, we can handle novel JCVs not in the dictionary.

5 Evaluation

The JCVs in Shin Meikai Kokugo dictionary were selected for the evaluation of our rules, because the meanings of JCVs can be judged objectively from their definition. Before the evaluation procedure, we categorized JCVs from the dictionary into idiomatic, fused, high frequency and exception categories.

Idiomatic JCVs are those where the meaning of the compound cannot be construed from the meaning of two verbs independently. The meaning of fused JCVs and high frequency JCVs can be inferred from each constituent. Fused JCVs are those which are used only in a specific context. High frequency JCVs can be divided into two verbs semantically and the case particle 'te' or 'de' can be inserted between the two verbs in certain cases. Exceptional JCVs are those with certain V2s such as *hajimeru* "start" and *tsuzukeru* "continue" which can be processed easily using only the definition of V2.

Since idiomatic and fused JCVs cannot be processed by our method, registering such JCVs in the dictionary is a reasonable approach for computer implementation. Exceptions may also be registered in the dictionary. However, all high frequency JCVs can be treated with our method and are designated as target words for evaluation.

5.1 Evaluation by using Japanese Dictionary

We extracted all JCVs which included any of 10 ambiguous V2s from the Japanese dictionary. Our target words for evaluation were all high frequency JCVs. In order to classify target words by semantic cluster, their dictionary definitions must include words related to some semantic cluster, such as *owaru* "finish" in the case of the **aspectual cluster**, *ue* "up" in the **spatial cluster** and *kurikaeshi* "again" in the **adverbial cluster**, etc.

Table 4 indicates the result of analyzing these JCVs. Idiomatic and fused JCVs and the name of semantic cluster are abbreviated in table 4, for example the **aspectual cluster** is shown as "ASPECT", etc. Half of the JCVs in the dictionary are regarded as idiomatic and fused words.

V2	IDIOM& FUSED	Target JCVs		
		ASPECT	SPACE	ADVERBI
agaru "go up"	15	3	20	8
ageru "lift"	41	5	21	2
iru "enter"	12		5	11
otosu "drop"	7		11	7
kaesu "send back"	9		16	10
kaeru "go back"	8		5	10
kakaru "hang onto"	7		11	
kakeru "hang"	18	3	14	
komu "insert"	95		58	16
sugiru "go past"	3		1	5
Total	215	11	162	69

Table 4. The list of JCVs in the dictionary

We took the following steps for evaluation.

(1) Extract target JCVs for evaluation from Japanese dictionary.
(2) Classify JCVs into each semantic cluster by referring to their definition.
(3) Assign the semantic features of Ruigo Shin Jiten to the V1 of JCVs. In the case that syntactic information is needed, it can be extracted from the examples of the dictionary.
(4) Prepare test sets including the target JCV, the semantic feature of V1 and the semantic cluster.
(5) Compare the test sets with our rules.
(6) Evaluate the accuracy of our rules.

We evaluated 242 JCVs from the dictionary, and obtained 211 correct rules and 31 errors. This corresponds to a high accuracy rate of 87.19%.

5.2 Discussion

As a result of the evaluation, 31 errors are observed. These errors can be divided into three types, corresponding to the lack of rules, the problem of semantic features and exceptions. Lack of rules including the semantic features of V1 and V2 has not yet been registered in our rules.

The second problem occurs where semantic features cannot be assigned to V1 appropriately. For instance, *koneru* "knead" is assigned *hendo* "fluctuation" as the semantic feature of V1, but the verb means motion for making something. The difference between *hendo* "fluctuation" and *seisan* "production" is important in identifying the semantic cluster, because *kone-ageru* "complete kneading" is in the **aspectual cluster**, but *maki-ageru* "roll up" is assigned the semantic feature of "fluctuation" in the **spatial cluster**. We consider that such verbs should be rearranged in an appropriate framework.

The errors classified as exceptions are those where an unusual usage of V1 causes the wrong cluster to be selected by our rules. For example, *moeru* "burn" used in *moe-agaru* "flare up" is assigned *bussho* "physical phenomena" as its semantic feature. *Moe-agaru* "flare up" should be regarded as **spatial cluster,** because of its dictionary definition that something burns with rising flames. However, a JCV with a V1 of *bussho* "physical phenomena" and V2 of *agaru* "go up" is classified into the **aspectual cluster** by our rules, similarly to *waki-agaru* "boil up" and *atatame-ageru* "finish heating". *Moe-agaru* "flare up" should be registered in the dictionary as an exception.

The accuracy of our rules is improved up to nearly 99% by the addition of 10 rules for 19 verbs and by rearranging the semantic features of 8 verbs. The result confirms the advantage of our method for disambiguating JCVs.

6 Conclusions and Future Work

We proposed a disambiguation method for JCVs based on disambiguation rules which make use of the semantic features of each constituent. We also clarified the characteristics of, and problems for, a treatment of JCVs. The obtained accuracy is 87.19% for our rules. This result shows the advantage of our method**.**

The disambiguation rules will be enhanced by analyzing the rest of the ambiguous V2 verbs shown in 2.3. This will complete our method for disambiguation. We will try to apply our method to

a machine translation system, comparing the results of our study with a study of phrasal verbs in English.

Reference

Timothy Baldwin and Francis Bond. 2002. Multiword Expressions: Some Problems for Japanese NLP, In *Eighth Annual Meeting of the Association of Natural Language Processing,* Keihanna, Japan, pages 379-382.

Masako Himeno. 1999. *The Structure of Compound Verbs and Semantic Usage*. Hitsuji Shobo, Tokyo.

Masako Himeno. 2001. The nature of Compound Verbs. Nihongogaku, 240(20):6-15.

Satoru Ikehara, et al. 1999. *A Japanese Lexicon*. Iwanami Shoten, Tokyo.

IPA, 1987. *Japanese basic verb dictionary IPAL.*

Taro Kageyama. 1993. *Grammar and Word Formation*. Hitsuji Shobo, Tokyo.

Taro Kageyama. 1999. Word Formation. The Handbook of Japanese Linguistics. Blackwell Publishers, Massachusetts, USA.

Kyosuke Kindaichi. 1999. *Shin Meikai Kokugo Dictionary*. 5th edition. Sanseido, Tokyo.

Mainichi NewsPapers. 1993. Mainichi Shinbun sha.

Yo Matsumoto. 1996. *Complex Predicates in Japanese.* CSLI Publications & Kurosio Publishers, Stanford & Tokyo.

Yo Matsumoto. 1998. The combinatory possibilities in Japanese V-V lexical compounds. *Gengo Kenkyu*, The Linguistic Society of Japan, 114:37-83.

Yuji Matsumoto, et al. 2000. *Morphological Analysis System Chasen Version2.2.1. Users Manual.*

Kazuaki Niimi, Youichi Yamaura, and Tokuko Utsuno. 1987. *Compound Verb*. Aratake Shuppan, Tokyo.

Susumu Oono and Masato Hamanishi. 1989. *Ruigo Shin Jiten*. Kadokawa Shoten, Tokyo.

Ivan A. Sag, Timothy Baldwin, Francis Bond, Ann Copestake, and Dan Flickinger. 2002. Multiword expressions: A pain in the neck for NLP. In *Proceedings of the Third International Conference on Intelligent Text Processing and Computational Linguistics: CICLING-2002*, Mexico City, Mexico, pages 1-15.

Satoshi Shirai, Yoshifumi Ooyama, Shinobu Takechi, Keiko Wakebe, and Hiroshi Aizawa. 1998. Compiling Japanese and English corpus for compound verbs of Japanese origin**.** In *57th Annual Meeting of IPSJ*, Nagoya, Japan, pages 267-268.

Yoshiko Tagashira. 1986. *Handbook of Japanese Compound Verbs.* Hokuseido, Tokyo.

Aline Villavicencio and Ann Copestake. 2002. Verb-particle constructions in a computational grammar of English. In *Ninth International Conference on Head-Driven Phrase Structure Grammar,* Seoul, South Korea.