Methods for Analyzing Japanese Predicate Idioms and Function Verb Expressions

MASAHIRO OKU*

The same expression in different sentences may take different meanings—a literal meaning or an idiomatic meaning—depending on its location or its modification by other words in a given sentence. Therefore, it is necessary to first determine whether the input sentences express literal meaning or whether they are being used in an idiomatic manner in order to perform analysis correctly. This paper proposes methods for analyzing predicate idioms and function verb expressions in Japanese sentences. These methods regard the characteristics of these expressions as the syntactic and semantic constraints required for the candidates to function as the expressions with idiomatic meanings. Only when expressions within the input sentence satisfy such constraints, are they analyzed as idiomatic expressions. Analyzing predicate idioms is performed by making the correspondence between the valency elements within the valency structure prepared for each predicate idiom and the case elements within the input sentence. A function verb expression consists of a verbal noun and a function verb. The meaning of the verbal noun is responsible for the whole meaning of the function verb expression and the function verb only has syntactic functions such as aspect or voice. Based on this characteristic, analyzing function verb expressions is performed on sentences with the verb form of such verbal nouns. These methods enable us to determine whether an expression within a sentence is being used according to its literal or idiomatic meaning. Note that a formal description of these constraints is not included in the general word dictionary. Through this separation, it is possible to describe information about idiomatic expressions without upsetting the description of words that are not idiomatic expressions. This, in turn, enhances dictionary maintenance.

1. Introduction

The same expression in different sentences may take different meanings—a literal meaning or an idiomatic meaning—depending on its location or its modification by other words in a given sentence. Therefore, it is necessary to first determine whether the input sentence is expressing literal meaning or whether it is being used in an idiomatic manner in order to perform analysis correctly.

A couple of methods have been suggested to analyze Japanese idiomatic expressions. One method (Method-1) registers an idiomatic expression as one lexical item in a dictionary. The other method (Method-2) provides information regarding idiomatic expressions in dictionary descriptions regarding the main words of the idiomatic expression [1-3]. However, neither method can distinguish whether the words are being used literally or idiomatically. The main difficulty lies in the idiomatic expressions that function as predicates which can be separated or modified. Separation or modifica-

tion can change the meaning. Neither Method-1 nor Method-2 is capable of coping with this problem. Further, with Method-2, the dictionary description related to the main words of idiomatic expressions tends to be longer and/or more complicated than the descriptions of words not used idiomatically. This imposes additional loads on the dictionary and raises problems regarding maintenance of the dictionary.

To solve these problems, we propose methods for analyzing predicate idioms and function verb expressions so that they can be separately processed from the general non-idiomatic expressions. The proposed methods regard the characteristics of these expressions as the syntactic and semantic constraints required for the candidate to take an idiomatic meaning in the sentence. In this way, it is possible to distinguish whether expressions appearing in sentences are being used literally or idiomatically. Furthermore, the description of these constraints is not included in the general word dictionary. Through this separation, it is possible to describe information about idiomatic expressions without upsetting the description of words that are not idiomatic expressions. This, in turn, contributes to better dictionary maintenance.

2. Classification of Idiomatic Expressions

According to [4], idiomatic expressions are defined as:

"Expressions in which the meanings of phrases or the meanings of elements constituting phrases have resulted in a divergence of one kind or another".

Idiomatic expressions thus defined are to be positioned and classified in the Japanese expressions as shown in Fig. 1 [4–8].

• General Phrases

These are generally used expressions that have been conventionally analyzed.

• Proverbs

These are expressions corresponding to short well-known sayings with historical or social sense of values [5].

• Predicate Idioms

YOUGEN (a verb, adjective, or adjective verb) forms the core of the phrase, and the idiom functions as a predicate within the sentence [5–7].

• Non-predicate Idioms

Idioms other than predicate idioms [5]. There are noun equivalent, e.g. "YAMA-NO KAMI" (literally "god of the mountains" but the idiom for "a nagging wife") or adverb equivalent, e.g. "TSURU-NO HITOKOE-DE" (literally "by one voice coming from a stork" but the idiom for "by the voice of authority").

• Function Verb Expressions

A function verb expression consists of a verbal noun and a function verb [4, 8, 9]. The meaning of the verbal noun is responsible for the whole meaning of the function verb expression and the function verb[1] only has syntactic functions such as aspect or voice.

In analyzing the above idiomatic expressions, proverbs should be dealt with as one lexical item in the dictionary since they are fixed phrases. It would be meaningless to divide them into structural elements for sentence analysis. The same can be said for non-predicate idioms as well. In contrast, both predicate idioms and function verb expressions function as predicates governing some case elements within the sentence, such as the "GA" case (the case element with postposition "GA") and the "WO" case (the case element with postposition "WO"). Thus, in order to analyze sentences including such expressions, the relationship between them and the case elements must be determined. Both predicate idioms and function verb expressions are, therefore, more interesting from a theoretical point of view.

Of the common idioms listed in [5], about 80% are predicate idioms and are structured syntactically as in the following Form (1);

Fig. 1   Classification of idiomatic expressions.

Fig. 2   Flow diagram of Japanese sentence analysis.

$$(MEISHI + JOSHI) + (YOUGEN + FUZOKUJI) \quad (1)$$

• MEISHI—noun.
• JOSHI—postposition.
• YOUGEN—verb, adjective, or adjective verb.
• FUZOKUJI—auxiliary verb or inflectional suffix.

Further, function verb expressions also have the structure given in Form (1). Thus, this paper assumes predicate idioms and function verb expressions to have Form (1) as they constitute the major portion of idiomatic expressions requiring structural analysis.

Hereafter, predicate idioms and function verb expressions with the structure of Form (1) will be termed idiomatic expressions.

3. The Position of Analysis for Idiomatic Expressions in Japanese Sentence Analysis

An outline of Japanese sentence analysis is shown in Fig. 2.

First, morphological and dependency analysis are performed for the input sentence. At this stage, consider-

---

[1]Such verbs are also called light verbs [9], delexical verbs [12], or functional verbs [13]. In this paper, such verbs are referred as "function verbs" according to [8], in which the characteristics of such verbs in modern Japanese are described.
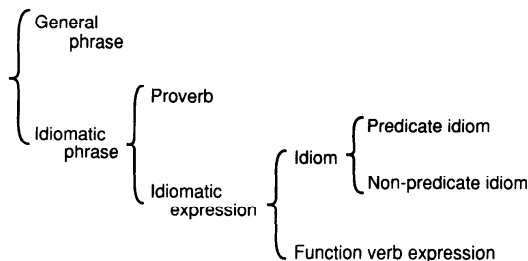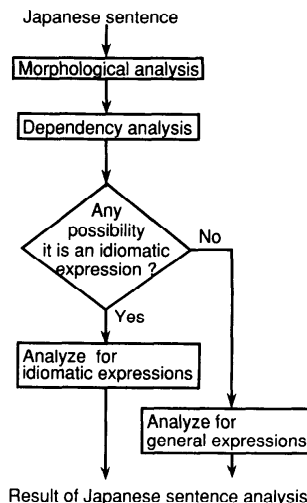
ing its ambiguity, several analytical results are obtained. Next, by using the valency structure, and by retaining only those analysis results which satisfy both the *JOSHI* and semantic constraints for the case elements of *YOUGEN*, the ambiguity is reduced [10].

Analysis of idiomatic expressions is utilized for reducing the ambiguity. The dependency between structural elements within an idiomatic expression is strong so that once an idiomatic expression within an input sentence can be identified, ambiguity may be readily reduced by adopting the interpretation.

An idiomatic expression has its structural elements co-occurring within a sentence (First condition), and must satisfy constraints required for it to have an idiomatic meaning (Second condition). The first condition relies on the existence of structural words and can be evaluated from the analysis results of dependency between complements and a predicate. Evaluation of the second condition needs to consider the constraints inherent in each idiomatic expression. The following discussions assume that morphological analysis and dependency analysis have already been concluded.

4. Analysis of Idiomatic Expressions

As mentioned in Chapter 2, regarding predicate idioms and function verb expressions, there is a need to analyze what kind of case elements they govern within the sentence. From the viewpoint of relationship between the meanings of the whole expression and of the structural elements, predicate idioms express the whole meaning by having all structural elements co-occurring. In contrast, function verb expressions differ in that the verbal nouns that are the structural elements are responsible for the major portion of the whole meaning.

4.1 Analysis of Predicate Idioms

4.1.1 Characteristics of Predicate Idioms

Predicate idioms are required to have their structural elements co-occurring and, therefore, the technique of handling these as one lexical item in a dictionary at the stage of morphological analysis can be considered. With this method, however, it is not possible to distinguish whether the expression within a sentence is meant to be a predicate idiom or whether it is to be interpreted literally. Furthermore, it is not possible to deal with the following characteristics which are peculiar to predicate idioms.

(a) Some predicate idioms are separable, which may result in different meanings.

Both Ex.a-1 and Ex.a-2 below include the expression "*KOKORO-GA SHIZUM-U*"—literally "mind sinks" but the idiom for "to be depressed"—and both have an idiomatic meaning despite the insertion of an additional word. In contrast, although both Ex.a-3 and Ex.a-4 include the expression "*ABURA-WO UR-U*"—liteally

"to sell oil" but the idiom for "to idle away one's time"—, Ex.a-3 has an idiomatic meaning while Ex.a-4 has a literal meaning depending on the word inserted between "*ABURA-WO*" and "*UR-U*".

Ex.a-1)	*KOKORO-GA* mind-SUBJ (to be depressed).		*SHIZUM-U* sink
Ex.a-2)	*KOKORO-GA* mind-SUBJ (to be depressed suddenly).	*KYUUNI* suddenly	*SHIZUM-U* sink
Ex.a-3)	*ABURA-WO* oil-OBJ (to idle away one's time).		*UR-U* sell
Ex.a-4)	*ABURA-WO* oil-OBJ (to sell oil cheaply).	*YASUKU* cheaply	*UR-U* sell

(b) The noun component of some predicate idioms may be modified by other words.

Although both Ex.b-1 and Ex.b-2 include the expression "*HANA-WO MOT-A-SERU*"—literally "to make one have flowers" but the idiom for "to save one's face"—, Ex.b-1 has an idiomatic meaning while Ex.b-2 has a literal meaning, dependent on the word modifying the noun component "*HANA-WO*". Both Ex.b-3, Ex.b-4 and Ex.b-5 include the expression "*SHIPPO-WO TSUKAM-U*"—literally "to grasp a tail" but the idiom for "to find one's weak point". Although Ex.b-4 and Ex.b-5 include the word modifying the noun component "*SHIPPO-WO*", Ex.b-4 has an idiomatic meaning while Ex.b-5 does not. Whether the meaning is idiomatic or not depends on semantic features of the modifying word.

Ex.b-1)	(to save one's face).	*HANA-WO* flower-OBJ	*MOT-A-SERU* have-causative
Ex.b-2)	*HARU-NO* spring's (to make one have the spring's flower).	*HANA-WO* flower-OBJ	*MOT-A-SERU* have-causative
Ex.b-3)	(to find one's weak point).	*SHIPPO-WO* tail-OBJ	*TSUKAM-U* grasp
Ex.b-4)	*KARE-NO* his (to find his weak point).	*SHIPPO-WO* tail-OBJ	*TSUKAM-U* grasp
Ex.b-5)	*NEKO-NO* cat's (to grasp the cat's tail).	*SHIPPO-WO* tail-OBJ	*TSUKAM-U* grasp

(c) Some predicate idioms comprise syntactic functions such as aspect or voice, and lose idiomatic meaning when such syntactic functions are excluded.

The expression "*SHUDAN-WO ERAB-U*" is included in both Ex.c-1 and Ex.c-2. This expression has no idiomatic meaning in itself. The expression "*SHUDAN-WO ERAB-A-NAI*"—literally "not select the means" but the idiom for "would stop at nothing"—consists of "*SHUDAN-WO ERAB-U*" and negative ending "*NAI*". Thus, Ex.c-1 has an idiomatic

meaning while Ex.c-2 does not.

Ex.c-1)   *SHUDAN-WO*   *ERAB-ANAI*
          means-OBJ   select-negative
          (would stop at nothing).

Ex.c-2)   *SHUDAN-WO*   *ERAB-U*
          means-OBJ   select
          (to select the means).

The characteristics of each of the foregoing vary with each individual predicate idiom and the expressions can be grasped as constraints enabling them to have meanings as predicate idioms. In other words, for expressions within sentences to have meanings as predicate idioms, structural elements not only must co-occur, but must also satisfy the constraints of each expression.

4.1.2   A Method for Analyzing Predicate Idioms

When analyzing predicate idioms, a valency structure is set up for the predicate idiom in the same manner as with *YOUGEN* [10]. Analysis is performed by finding correspondences between the valency structure and the case elements within the input sentence.

As mentioned in the preceding section, it is not possible to handle a predicate idiom as a single lexical item. Thus, in this paper, analysis by the use of valency structure of predicate idioms is attempted only in instances where analysis results indicate linkage of structural elements in a state of dependency and only when the required constraints are satisfied. This method requires description of the information expressing constraints required for a predicate idiom together with the valency structure of the predicate idiom.

The valency structure of a predicate idiom needs to be referred to only when satisfying constraints so that information expressing constraints may be described separately from the valency structure. There is a need to describe the valency structure of predicate idioms within the same framework as *YOUGEN* to allow their common processing in the stages following analytical processing.

Idiom control table

Candidate of predicate idiom	Required constraints			Pointer
	Possibility of insertion	Possibility of modification	Syntactic function	

Idiom dictionary

Predicate idiom		
N1	SC:	JC:
N2	SC:	JC:
:	:	:
Ni	SC:	JC:

(Legend)
Ni   ··· Label of valency element
SC   ··· Semantic constraints of a head noun
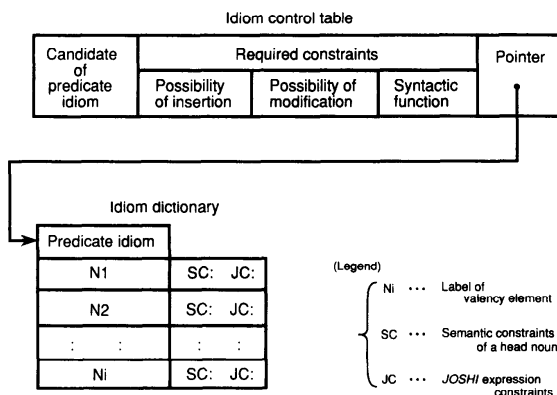JC   ··· *JOSHI* expression constraints

Fig. 3   Structure of the idiom control table and the idiom dictionary.

From the foregoing, an idiom control table, which expresses information describing constraints, and an idiom dictionary describing idiom valency structure as shown in Fig. 3, are prepared.

The idiom control table and the idiom dictionary are connected by pointers. Therefore, when constraints are satisfied, the idiom valency structure is readily obtained by following the appropriate pointer.

The idiom control table contains the information needed to enable handling of characteristics (a) through (c) (refer to section 4.1.1). Specifically, the following four points are described.

- Possibility of insertion (pro and con of insertion of other elements between structural elements of predicate idioms).
- Possibility of modification (pro and con of modifying a noun component constituting a predicate idiom. If pro, the label of valency element adopted by the modifying element, the semantic constraints of a head noun and *JOSHI* expression constraints are described).
- Syntactic function (such as causative and negative that must be comprised in a predicate idiom).
- Pointers for an idiom dictionary.

For the idiom dictionary, the following two constraints must be described for each valency element in the same manner as in the dictionary of the valency structure for *YOUGEN*.

- Semantic constraints of each head noun.
- *JOSHI* expression constraints.

An outline of the analysis of predicate idioms is shown in Fig. 4, and an actual analysis example in Fig. 5. Predicate idiom analysis is invoked by idiom candidates (the co-occurrence of certain structural elements of a predicate idiom). In actual practice, the noun and *YOUGEN* which are found to be structural elements of predicate idioms are marked in the word dictionary used in morphological analysis.[1] Elements having a dependency relation with the marked noun and *YOUGEN* are extracted as candidates for a predicate idiom.

The example in Fig. 5 shows the two marked words, "*HANA*" (flowers) and "*MOT-U*" (have). With the two words linked in a dependency relation, they are extracted as a candidate "*HANA-WO MOT-U*" (have flowers), which is the expression excluding causative function "*SERU*" from the predicate idiom "*HANA-WO MOT-A-SERU*" (to save one's face).

When a candidate is extracted and satisfies the constraints described in the idiom control table, the idiom

---

[1]With this method, all of the words in the word dictionary need to be marked as to whether they can be candidates for structural elements of predicate idioms. If it is a method of selecting candidates from a combination of a noun and *YOUGEN* having a dependency relation, such descriptions are not necessary. It is possible to separate information related to predicate idioms completely from information related to general expressions. However, this method involves meaningless searches and is inefficient and impractical.
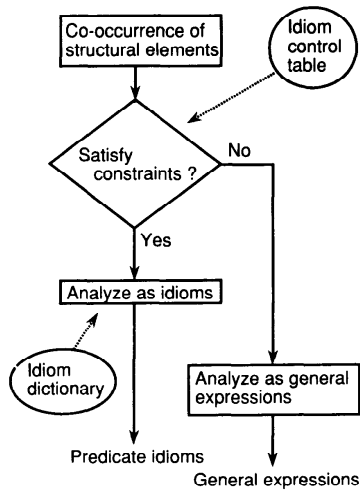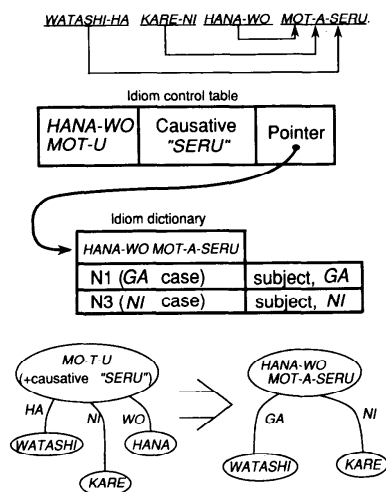
Fig. 4 Analysis flow of predicate idioms.

Fig. 5 An example of idiom analysis.

dictionary is referred to and analysis for a predicate idiom is performed. In the example in Fig. 5, *"HANA-WO MOT-U"* is extracted as a candidate. The input sentence has satisfied the constraint (syntactic function=causative *"SERU"*) regarding *"HANA-WO MOT-U"* in the idiom control table. Thus, analysis is performed using the valency structure in the idiom dictionary resulting in the analysis output shown in the figure.

When constraints described in the idiom control table or in the idiom dictionary are not satisfied, analysis is performed using the valency structure of *YOUGEN* as a general expression.

4.2 Analysis of Function Verb Expressions

4.2.1 Characteristics of Function Verb Expressions

A function verb expression consists of a verbal noun and a function verb. The meaning of the verbal noun is responsible for the whole meaning of the function verb expression and the function verb only has syntactic functions. Due to this characteristic, to attempt to handle it in the same manner as *YOUGEN* in structural analysis would pose the following problems.

(a) Some function verb expressions have proper case relations. The case relation is identical with the case relation yielded by the verb form of the verbal noun that consists of the function verb expression.

(b) Some function verbs provide syntactic functions, e.g. aspect or voice.

Further, as a method of analyzing function verb expressions, the method of using the valency structure of each function verb expression, as in the case of predicate idioms, can be considered. However, with this method, a number of valency structures must be prepared to match the number of function verb expressions. Function verb expressions are expressions combining verbal nouns with function verbs so that their numbers would be the multiple of the number of verbal nouns and of function verbs. This number is much larger than the number of predicate idioms. Therefore, the method that involves describing a valency structure for each function verb expression is not practical.

4.2.2 A Method for Analyzing Function Verb Expressions

To solve the foregoing problems, this paper proposes a method for analyzing function verb expressions. The method, shown in Fig. 6, handles a function verb expression as an expression with a predicate for the verb form of the verbal noun.

The method is based on the characteristic of "the meaning of the verbal noun is responsible for the whole meaning of the function verb expression and the function verb only has syntactic functions". This method has several advantages in that it is capable of reducing the number of analytical rules to practically the same number as there are function verbs. In analyzing a sentence including a function verb expression based on this method, the following five points must be taken into account.

(a) *JOSHI* expressions must be converted.

For example, in Ex.a-1 in Fig. 6(a), postposition *"NI"* must be converted to postposition *"WO"* and, in Ex.b-1, the conversion of postposition *"NO"* to postposition *"NI"* must be required depending on raising the case elements.

(b) Syntactic function provided by the function verb must be added.

In Ex.b-1 in Fig. 6(a), for example, since the function verb *"UKE-RU"* has a syntactic function "passive",
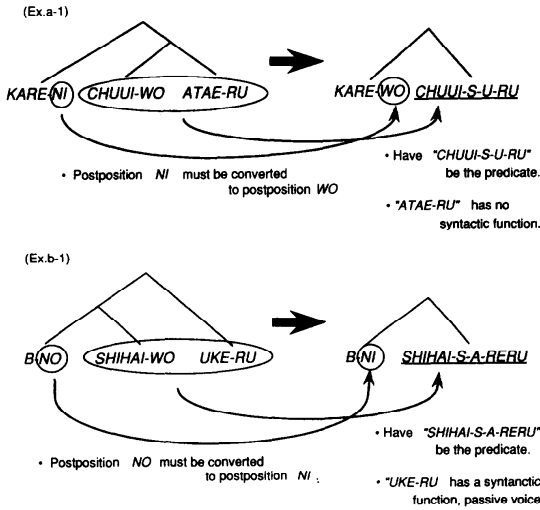
(Ex.a-1)

- Postposition *NI* must be converted
  to postposition *WO*

- Have *"CHUUI-S-U-RU"*
  be the predicate.

- *"ATAE-RU"* has no
  syntactic function.

(Ex.b-1)

- Postposition *NO* must be converted
  to postposition *NI* .

- Have *"SHIHAI-S-A-RERU"*
  be the predicate.

- *"UKE-RU"* has a syntactic
  function, passive voice.

Fig. 6(a)   Characteristics of function verb expressions.

- Usually applied rule for function verb *"ATSUME-RU"*

Verbal noun + function verb *"ATSUME-RU"* ➡ Add to syntactic function,
"**passive**".

(Ex.e-1)

  - Case obeying the rule :

    CHUUMOKU-WO   ATSUME-RU ➡   CHUUMOKU-S-A-RERU
    (attention   (gather)   (attention-passive)
    -DOBJ)

(Ex.e-2)

  - Case that *"ATSUME-RU"* is used as a main verb :

    SHOMEI-WO   ATSUME-RU ➡   SHOMEI-S-A-RERU
    (signature   (gather)   (sign-passive)
    -DOBJ)

                                         SHOMEI-WO   ATSUME-RU
                                         "ATSUME-RU" is used as a main verb.

Fig. 6(b)   Characteristics of function verb expressions.

- Usually applied rule for function verb *"ATAE-RU"*

Intransitive verbal noun + function verb *"ATAE-RU"* ➡ Add to syntactic function,
"**causative**".

(Ex.d-1)

  - Case obeying the rule :

    KANDOU-WO   ATAE-RU ➡   KANDOU-S-A-SERU
    (impression   (give)   (be-impressed-causative)
    -DOBJ)

(Ex.d-2)

  - Exception :

    EIKYOU-WO   ATAE-RU ➡   EIKYOU-S-U-RU
    (effect-DOBJ)   (give)   (effect (no syntactic function))
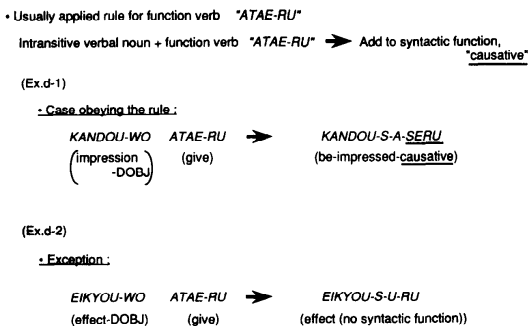
Fig. 6(c)   Characteristics of function verb expressions.

this function must be added to the expression with a new predicate (the verb form of the verbal noun).

(c)   The same function verbs provide different syntactic functions depending on the characteristics of the verbal nouns.

For example, the function verb *"UKE-RU"* co-occurring with the transitive verbal noun,   shown in Ex.c-1, provides the syntactic function "passive" while *"UKE-RU"* co-occurring with the intransitive verbal noun, shown in Ex.c-2, has no syntactic function.

Ex.c-1)   *SHIHAI-WO*   *UKE-RU*   —*SHIHAI-S-A-RERU*
              control-DOBJ   receive   control-passive

Ex.c-2)   *KANMEI-WO*   *UKE-RU*   —*KANMEI-S-U-RU*
              impression-DOBJ   receive   be-impressed
                                                              (no syntactic function)

(d)   Discrete analytical rules are needed to handle some exceptional combinations of function verbs and verbal nouns.

Although both Ex.d-1 and Ex.d-2, shown in Fig. 6(b), include the function verb *"ATAE-RU"* and intransitive verbal nouns, Ex.d-1 obeys the usual rules for function verb *"ATAE-RU"* while Ex.d-2 requires the use of a discrete rule.

(e)   Certain verbs function as the function verbs or the main verbs depending on the co-occurring verbal nouns.

For example, the verbal noun *"CHUUMOKU"*, in Ex.e-1 in Fig. 6(c), allows the verb *"ATSUME-RU"* to work as the function verb while the verbal noun *"SHOMEI"* in Ex.e-2 allows it to work as the main verb.

Both characteristics (a) and (b) can be interpreted as the rules for analyzing an expression which handle the input sentence as the predicate for the verb form of the verbal noun. Characteristic (c) can be regarded as a precondition to decide analytical rules. Both characteristics (d) and (e) require individual processing.

Therefore, both analytical rules and constraints will be described in the function verb expression control table, and information regarding individual processing will be described in the individual control table.

In order to deal with the characteristics (a) through (c) mentioned above, the following five points will be described in the function verb expression control table.

- Description of transitive/intransitive of the verb form of the verbal noun (constraint).
- Semantic features of verbal nouns (constraint).
- Syntactic function provided by function verbs (analytical rule).
- Rules for *JOSHI* conversion (analytical rule).
- Pointers for individual control tables (for the purpose of individual processing).

Individual control tables are to be prepared for individual function verbs. Each individual control table is to describe the following three points for each verbal noun to enable dealing with both characteristics (d) and (e).
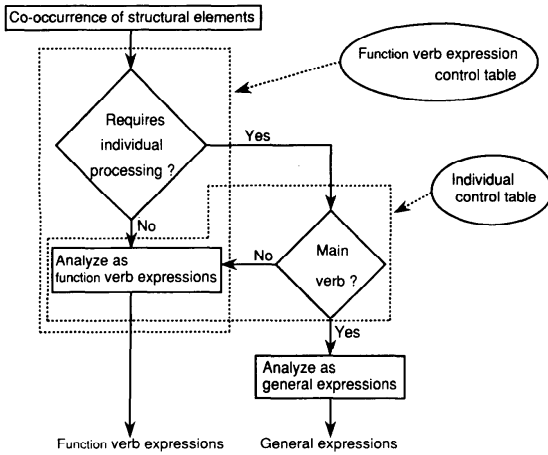
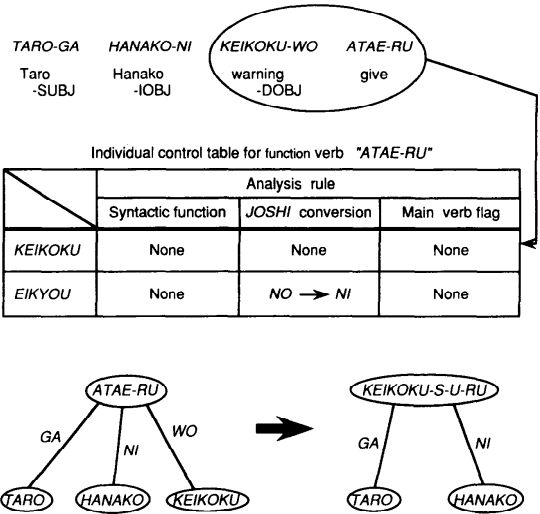Fig. 7   Analysis flow of function verb expressions.

Function verb expression control table

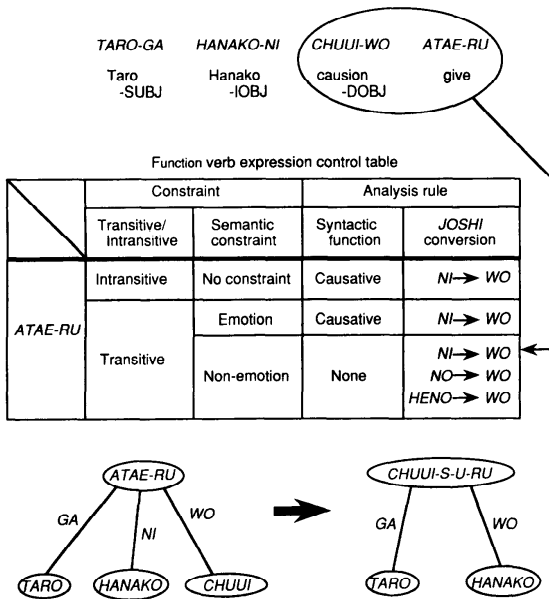|	Constraint		Analysis rule	
	Transitive/ Intransitive	Semantic constraint	Syntactic function	*JOSHI* conversion
	Intransitive	No constraint	Causative	NI→ WO
ATAE-RU		Emotion	Causative	NI→ WO
	Transitive	Non-emotion	None	NI→ WO NO→ WO HENO→ WO

Fig. 8   An example of the analysis of a function verb expression with the function verb expression control table.

- Syntactic function provided by the function verb.
- Rules for *JOSHI* conversion.
- Flags to indicate whether it is used as a main verb or not.

An outline of the analysis of function verb expressions is shown in Fig. 7 and actual examples in Fig. 8 and Fig. 9.

The analysis of function verb expressions is invoked by the co-occurrence of verbal nouns and function verbs. In actual practice, function verbs are marked in the word dictionary used in morphological analysis. The analysis is activated only when the verbal noun

Individual control table for function verb "*ATAE-RU*"

	Analysis rule		
	Syntactic function	*JOSHI* conversion	Main verb flag
KEIKOKU	None	None	None
EIKYOU	None	NO → NI	None

Fig. 9   An example of the analysis of a function verb expression with the individual control table.

modifies a verb so marked. In the examples in Fig. 8 and 9, the word "*ATAE-RU*" is marked to indicate that it is a function verb. Both "*CHUUI*" and "*KEIKOKU*" are verbal nouns so that both "*TARO-GA HANAKO-NI CHUUI-WO ATAE-RU*" and "*TARO-GA HANAKO-NI KEIKOKU-WO ATAE-RU*" undergo function verb expression analysis.

In the analysis of function verb expressions, applicable rules are obtained through searching the function verb expression control table. When the verbal noun does not require individual processing (i.e. when the verbal noun is not listed in the individual control table), analysis is performed by regarding the verb form of the verbal noun as the predicate as described in the function verb expression control table. When the verbal noun requires individual processing (i.e. when the verbal noun is listed in the individual control table), analysis is performed according to the applicable description in the individual control table. With the example in Fig. 8, the verbal noun "*CHUUI*" does not require individual processing. Moreover, the semantic feature of the verbal noun "*CHUUI*" is not "emotion" and its verb form "*CHUUI-S-U-RU*" is transitive. Therefore, *JOSHI* conversion "from *NI* to *WO*" described in the function verb expression table is performed, and as a result of analysis, "*TARO-GA HANAKO-WO CHUUI-S-U-RU*" (Taro cautions Hanako) is obtained. In contrast, the verbal noun "*KEIKOKU*" requires individual processing. Thus, analysis is performed according to the description in the individual control table, and as a result, "*TARO-GA HANAKO-NI KEIKOKU-S-U-RU*" (Taro warns Hanako) is obtained.
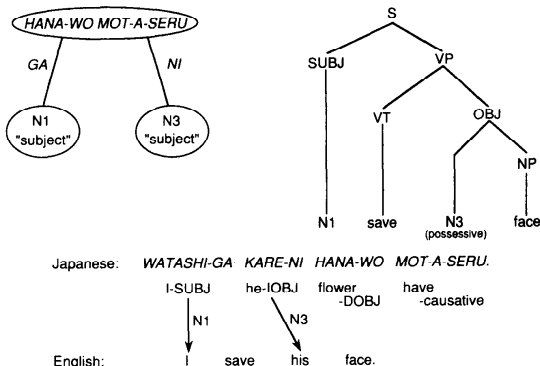
Fig. 10   An example of translating a Japanese predicate idiom into English.

5. Application to a Japanese to English Translation System

We have been developing a Japanese to English machine translation system based on a multi-level-translation method [11]. This method classifies expressions into three levels dependent on the strength of the linkage between the Japanese syntactic structures and their meanings. In this system, English syntactic structures corresponding to Japanese predicate idioms are prepared, as shown in Fig. 10, in order to complete the translation. English syntactic structures and Japanese valency structures for predicate idioms are connected through the labels of valency elements (N1 and N3 in Fig. 10). In this way, it becomes possible to correctly translate idiomatic expressions, since structural conversion rules become a collection of conversion rules peculiar to expressions.

There are Japanese expressions consisting of a complement and a predicate which correspond to a single word or an idiom in English (e.g. *SE-GA TAKA-I* (literally "back is high" but meaning is "tall"). These expressions cannot be translated correctly in a word to word translation. Expressions of this kind have the same characteristics as predicate idioms.

(a)   These expressions are separable.

For example, the Japanese expression "*KOURITSU-GA YO-I*" which corresponds to the single English word "efficient" still has the same meaning in both Ex.a-1 and Ex.a-2 regardless of the insertion of an additional word.

Ex.a-1)	*KOURITSU-GA*		*YO-I*
	efficiency-SUBJ		good
	(to be efficient).		
Ex.a-2)	*KOURITSU-GA*	*HIJOUNI*	*YO-I*
	efficiency-SUBJ	very	good
	(to be very efficient).		

(b)   The noun component may be modified.
An example is shown below.   "*SE-GA TAKA-I*"

which corresponds to the English equivalent "tall" is included in both Ex.b-1 and Ex.b-2. Since this expression allows the other word to modify the noun component, Ex.b-2 still has the same meaning.

Ex.b-1)	*SE-GA*		*TAKA-I*
	back-SUBJ		high
	(to be tall).		
Ex.b-2)	*KARE-NO*	*SE-GA*	*TAKA-I*
	his	back-SUBJ	high
	(he is tall).		

Thus, it is possible to analyze these in the same way as predicate idioms. This will overcome the different structure of Japanese and English to a certain extent and assure more accurate translations. Currently, including these expressions, about 3,000 predicate idioms have been formulated in the framework mentioned in section 4.1.2.

Regarding function verb expressions, we have mentioned previously in Section 4.2.1 that we cannot analyze function verbs as general verbs. Therefore, by analyzing function verb expressions accurately using the method proposed in section 4.2.2, it becomes possible to make an accurate translation of function verb expressions. Currently, 16 function verbs are formulated in the function verb expression table.

6. Conclusion

This paper has dealt with idiomatic expressions of the syntactic structure as (*MEISHI + JOSHI*) + (*YOUGEN + FUZOKUJI*). The proposed methods regard the characteristics of each expression as constraints required for idiomatic usage. Depending on whether these constraints have been satisfied or not, it is possible to distinguish whether expressions within the sentences have been used literally or idiomatically. Also, the methods enable separation of information related to idiomatic expressions from information related to words and thereby contribute to updating, deletion and other aspects of dictionary maintenance.

Examination of some 12,000 newspaper article sentences and some 9,000 sentences related to scientific and technological areas has revealed that about 2% are predicate idioms and about 3% are function verb expressions. The methods proposed in this paper were able to deal with about 90% of the idiomatic expressions encountered. Therefore, these methods can be expected to upgrade analytical precision in Japanese sentence analysis by approximately 4.5%.

Problems to be attacked in the future include expansion of the method for predicate idioms to cover predicate idioms consisting of two or more complements and a predicate (for example, "*KAO-NI DORO-WO NU-RU*"—literally to smear one's face with mud, but the idiom is "to embarrass") and to improve the precision in evaluating whether an expression is a predicate idiom or not by means of context informa-

tion. Regarding function verb expressions, there are occasions where verbal nouns constituting function verb expressions are modified by the sentence (e.g. *"NYUURYOKU-BUN-NI OUJI-TA KAISEKI-WO OKONA-U"*). In such cases, these must be analyzed as a compound sentence of two or more simple sentences linked by a conjunction (e.g. *"NYUURYOKU-BUN-NI OUJI-TE KAISEKI-S-U-RU"*). This cannot be handled within the framework described in section 4.2.2. Therefore, problems for analyzing function verb expressions to be solved in the future include developing a function verb expression control table so as to be capable of dealing with such situations.

Acknowledgment

References
1. NISHIDA, F. and TAKAMATSU, S. Intermediate expression transfer and sentence generation in machine translation, *J. IPS Japan*, **26**, 10 (1985), 1165–1173 (in Japanese).
2. NAGAO, M. and TSUJII, J. The transfer phrase of the Mu machine translation system, *Proc. of COLING 86* (1986), 97–103.
3. NAGAO, M. and TSUJII, J. Lexical and structural transfer in a machine translation system, *J. IPS Japan*, **26**, 11 (1985), 1261–1270 (in Japanese).
4. MURAKI, S. *KANYOUKU·KINOU DOUSHI KETSUGOU ·JIYUU-NA GO-KETSUGOU* (Japanese idioms, function verb expressions and general expressions, *Nihon-go-gaku* (Japanese Philology), **4** (1985), 15–27 (in Japanese).
5. Compiled by MIYAJI, H. *KANYOUKU-NO IMI-TO YOUHOU* (Meaning and usage of Japanese idioms), *Meiji-Shoin* (1982) (in Japanese).
6. MORITA, Y. *DOUSHI KANYOUKU* (Japanese verbal idioms), *Nihon-go-gaku* (Japanese Philology), **4** (1985), 37–43 (in Japanese).
7. NISHIO, T. *KEIYOUSHI KANYOUKU* (Japanese adjective idioms), *Nihon-go-gaku* (Japanese Philology), **4** (1985), 45–33 (in Japanese).
8. MURAKI, S. *NIHONGO-NO KINOU DOUSHI HYOUGEN-WO MEGU-TTE* (On function verbs in modern Japanese)'', *Kokuritsu Kokugo Kenkyuusho Houkoku 65, Kenkyuu houkoku-shuu* (2) (1980), 17–75 (in Japanese).
9. LIVE, A. H. The Take-Have Phrasal in English, *Linguistics*, **95** (1973), 31–50.
10. HAYASHI, Y. On the Japanese sentences analysis based on the valency structure, *Working group report on natural language processing, IPSJ*, 62–6 (1987) (in Japanese).
11. IKEHARA, S., MIYAZAKI, M., SHIRAI, S. and HAYASHI, Y. Speaker's recognition and multi-level-translating method based on it, *Trans. IPS Japan*, **28**, 12 (1987), 1269–1279 (in Japanese).
12. TANAKA, H. and AIZAWA, T. A method for translation of English delexical verb—deverbal noun phrases into Japanese, *Working group report on natural language processing, IPSJ*, 82–6 (1991) (in Japanese).
13. SUEMATSU, H., SUGIURA, M., ARIOKA, M. and JONES, T. Descriptive and representational frameworks for English collocations in an English dictionary, *Working group report on natural language processing, IPSJ* 86–3 (1991).