

# Completion of Japanese Sentences by Inferring Function Words from Content Words

Koji KAKIGAHARA and Teruaki AIZAWA

ATR Interpreting Telephony Research Laboratories  
Twin 21 Bldg. MID Tower, 2-1-61 Shiromi,  
Higashi-ku, Osaka 540 Japan

## Abstract

A method of generating a Japanese sentence by inferring function words from content words using valency patterns is presented.

A procedure for selecting an appropriate function word, on the assumption that correct content words have been selected for a given phrase lattice, is described. A method of inferring a correct verb when verbs are recognized less accurately than nouns by the speech recognition system is described. Sentences are produced from content words as inputs by using the valency patterns obtained from collected dialogue sentences in a restricted task domain. Using the semantic features of preceding nouns and valency patterns allow a fairly restricted number of candidate verbs to be inferred.

This method eliminates possible errors at the interface between speech recognition and machine translation (component technologies of an Automatic Telephone Interpretation system) and selects the most appropriate candidate from a lattice of typical phrases output by the speech recognition system.

## 1 Background and problems

An Automatic Telephone Interpretation system is a facility which enables a person speaking in one language to communicate readily by telephone with someone speaking another. Three constituent technologies are necessary for such a system: speech recognition, machine translation, and speech synthesis.

Basic research in each of these technologies to

develop an interpretation system between Japanese and English has already started. For this purpose, however, an effective interface between speech recognition and machine translation is vital because output from the speech recognition module will inevitably contain errors which the machine translation module cannot accept.

This paper proposes a method of generating a Japanese sentence by inferring function words from content words using valency patterns. This technique is aimed at the selection of the most appropriate candidate from a typical phrase lattice that may result from a speech recognition system.

## 2 Basic assumptions

In this study the following restrictions relevant to the interface problem are assumed:

- (1) A Japanese sentence usually consists of a certain number of noun phrases followed by a verb phrase at the end. The basic unit of speech recognition is assumed to be a continuously uttered phrase, so that any input to the machine translation module is a 'phrase lattice', i.e., a set of phrase candidates hypothesized by the speech recognition module.
- (2) The range of telephone conversation tasks is restricted to inquiries from a researcher to a clerk about an international conference concerning the main topic of the conference, deadlines for paper submission, exhibitions, social events, accommodation, payment, cancellation, etc..

utterance	phrase 1 'genkou-no'	phrase 2 'shimekiri-wa'	phrase 3 'itsu-desuka'
results of speech recognition	'genkou' 8 'kentou' 5 'nentou' 1	'shimekiri' 7 'shigeki' 3	'itsu' 5 'isu' 2 'iku' 1
	'wo' 4 'mo' 3 'no' 1	'wa' 3 'ga' 2	'desuka' 4 'deshita' 3 'bekika' 1
	'wo' 4 'mo' 3 'no' 1	'niwa' 4 'niga' 3	'desuka' 4 'deshita' 3 'bekika' 1
	'wo' 4 'mo' 3 'no' 1		'desuka' 4 'deshita' 3 'bekika' 1

Figure 1 An example of a phrase lattice.

### 3 Research goal

#### 3.1 A phrase lattice as the result of speech recognition

Consider a Japanese sentence consisting of two noun phrases and one verb phrase:

'genkou-no shimekiri-wa itsu-desuka'.  
(‘When is the deadline for a manuscript?’)

Usually a Japanese phrase begins with a jiritsugo-word (J-word for short) such as a noun or verb, and ends with a sequence of fuzokugo-words (F-words for short) such as postpositional particles or auxiliary verbs. In the above notation, boundaries between J-words and F-words are explicitly indicated by hyphens, and all F-words are italicized.

Figure 1 shows an example of a phrase lattice for this sentence obtained as the result of speech recognition.

Notice that there are candidates for both J-words and F-words together with a recognition score of the probability that the word is correct. The problem is to select the most appropriate candidate from this phrase lattice.

#### 3.2 Selection-by-generation approach

Attention is focused on candidates for F-words, assuming that J-words have already been correctly selected by a suitable method.

The assumption that J-words have been correctly selected is realistic if the task domain is limited enough to allow a high recognition rate for J-words and a knowledge-base, etc. is available for the limited task domain. Techniques related to this procedure are now being studied. Of the J-words, the predicate at the end of a sentence is less accurately recognized by the speech recognition module than nouns. A method to solve this problem will be discussed in the second half of this paper.

In Figure 1, for instance, it is assumed that a sequence of J-words:

'genkou'            'shimekiri'        'itsu'  
(‘manuscript’)    (‘deadline’)        (‘when’)

has been correctly selected according to the recognition scores. Corresponding to these J-words, there are three sets of candidates for F-words in the phrase lattice:

'wo'                'wa'                'desuka'  
'mo'                'ga'                'deshita'  
'no'                                     'bekika'.

Our major concern here is the subproblem of selecting the most appropriate one in each of these sets.

This sub-problem is characteristic of the Japanese language. In fact, as easily seen in the above example, frequently used F-words, specifically those indicating grammatical cases such as 'ga', 'wo', 'ni', etc., are too short to be recognized correctly. Their recognition scores are much lower than those of J-words. But in Japanese it is often possible to infer the meaning of a given sentence from the sequence of J-words when the task domain is narrow.

Our method of selecting the correct F-words is composed of two steps: 1) generate a meaningful sentence by inferring suitable F-words for a given sequence of J-words, and 2) compare these inferred F-words with the candidates in the phrase lattice to select those most appropriate.

This idea of 'selection-by-generation' distinguishes this approach from previous ones: Hayes et al. [1] for English or Niedermair [2] for German. In this paper only Step 1, which is considered the key step, will be discussed.

### 4 Generating a sentence by inferring F-words

The task domain is restricted to inquiries about an international conference, and therefore the dialogue is basically a repetition of simple questions and answers. This increases the probability of inferring the correct F-words for each phrase.

#### 4.1 Key information for the inference

The following types of information are used to infer F-words. The information is described in a lexicon of J-words.

##### (1) Semantic features of nouns and valency patterns

First, two types of semantic features are set up for nouns appearing in the restricted task domain. One is a general type of semantic feature, independent of the task domain, such as abstract, action, concrete, human, location, time, number and diversity. The other is a specific type of semantic feature dependent on the task domain. Table 1 shows examples of such features.

Using these semantic features valency patterns of the basic predicates necessary in the task domain are defined. As an example, the predicate 'okuru' ('send' in English) is given the following valency patterns:

N[con/-tra]'wo' + V,  
N[con/-tra]'wa' + N[loc]'ni' + V,  
N[con/-tra]'wa' + N[hum]'ni' + V,  
N[con/-tra]'wa' + N[tim/pro]'madeni' + V,  
N[tim/pro]'madeni' + N[con/-tra]'wo' + V,  
N[hum]'ni' + N[con/-tra]'wo' + V,  
N[hum]'ga' + N[con/-tra]'wo' + V,  
N[hum]'ga' + N[con/-tra]'wo' + N[hum/-pro]'ni' + V,  
N[hum/-pro]'wa' + N[con/-tra]'wo' + N[hum]'ni' + V,  
etc..

The first valency pattern in this list, for instance, specifies that the predicate V ('okuru') can take one noun phrase consisting of a noun with the general semantic feature 'concrete' / specific semantic feature non-'transport' and F-word 'wo'.

In this way the valency patterns summarize the basic J-word and F-word relationships, and thus give the most essential information for inferring F-words from a given sequence of J-words.

Table 1 Semantic features

general	specific	example
abstract	logic	'iron'(theory), 'houhou'(method)
	state	'yousu'(state), 'baai'(case)
	language	'nihongo'(Japanese), 'eigo'(English)
	learning	'bunya'(field), 'senmon'(specialty)
	intention	'kyoumi'(interest), 'kibou'(hope)
	value	'hitsuyou'(necessity)
	sign	'namae'(name)
	labor	'youken'(business)
concrete	document	'genkou'(manuscript), 'youshi'(form)
	transport	'basu'(bus), 'tikatetsu'(subway)
	article	'syashin'(photograph)
	money	'okane'(money), 'kado'(cash card)
	pronoun	'kore'(this), 'dore'(which)
human	human	'happyouya'(presenter)
	pronoun	'dare'(who)
location	institution	'kaijyou'(hall), 'hoteru'(hotel)
	region	'kyoto'(Kyoto), 'kaigai'(foreign)
	position	'temoto'(hand), 'aida'(between)
	pronoun	'doko'(where)
time	time	'jikan'(time), 'ima', (now) 'ato'(after)
	pronoun	'itsu', (when) 'nanji'(what time)
number	amount	'ninzuu'(the number of people)
	unit	'en'(yen), 'doru'(dollar)
	cost	'tourokuryou'(registration fee)
	price	'muryou'(free), 'ikura'(how much)
act		'sanka'(participation), 'yotei'(plan)
diverse		'nani'(what), 'hoka'(else)

These valency patterns are obtained from the valency patterns of predicates (obtained from dialogues collected for the task of inquiries about an international conference). If necessary, certain modifications such as omission of the nominative case, modification of the word sequence or of the F-words, and addition of interrogative pronouns are carried out. In dialogue sentences, the nominative case such as 'watashi' ('I') or 'anata' ('you') is seldom used; hence, the nominative case is usually not included in the valency patterns. To describe the modification of the two valency patterns for 'okuru' ('send'):

N[*tim/pro*] 'madeni' + N[*con/-tra*] 'wo' + V (a)

N[*con/-tra*] 'wa' + N[*tim/pro*] 'madeni' + V (b)

If the noun N[*con/-tra*] of the valency pattern (a) becomes the subject, the F-word is replaced with 'wa' and the word sequence is changed, often resulting in the valency pattern (b). Interrogative pronouns are added to produce valency patterns specific to interrogative sentences because a large number of questions occur in this task domain.

In a limited task domain, even individually optional cases behave in a similar way to the obligatory case for each predicate. Therefore, the optional cases are described in these valency patterns. When valency

patterns were prepared for 65 words working as predicates, an average of about 11 valency patterns were produced for each predicate. Details will be discussed in Chapter 5.

## (2) Connection of two nouns by F-word 'no'

It is inferred that nouns which cannot be processed through valency patterns are likely to be connected with the F-word 'no' (roughly corresponding to 'of' in English) in the form 'A no B', where A and B denote nouns. For a given noun A, the other noun B can also be specified through the semantic features. For instance, the noun 'kaigi' ('conference') can be joined with other nouns as follows:

'kaigi' + 'no' + N[*hum/hum*],  
 'kaigi' + 'no' + N[*abs*],  
 'kaigi' + 'no' + N[*tim/tim*],  
 'kaigi' + 'no' + N[*loc/ins,pos*],  
 N[*abs*] + 'no' + 'kaigi',  
 N[*num/amo*] + 'no' + 'kaigi', etc..

As shown above, whether or not to insert the F-word 'no' is automatically determined by presetting which nouns are to be connected with the F-word 'no'.

## (3) Syntactic information

Pure syntactic knowledge is also useful in this process. It is known that, in Japanese, no F-word can be attached to an adverb or a conjunction, and that a verb in conditional form can be connected with an adjective via F-words such as 'ba'.

In addition, the following rules are used, for example:

Continuative verb form + F-word + verb  
 → conjunctive particle 'te',  
 Conclusive verb form + F-word + verb  
 → conjunctive particle 'to',  
 Continuative verb form + F-word + adjective  
 → conjunctive particle 'te',  
 conjunctive particle 'temo',  
 Attributive verb form + F-word + noun  
 → no F-word, etc..

## 4.2 Outline of the process of inferring F-words

Figure 2 (a) illustrates a process of inferring F-words in a Japanese sentence:

'kaijyou-de kaigi-no yousu-wo  
 rokuonshi-temo ii-desuka',  
 (May I record the speech of the conference  
 at the hall?).

In this case it is assumed that J-words 'kaigi', 'yousu', 'rokuonshi', and 'ii' are correctly recognizable.

The inference proceeds as follows: 1) syntactic information can connect 'rokuonshi' and 'ii' with F-words 'te' or 'temo' and 'desuka' to generate the phrases 'rokuonshi-te' or 'rokuonshi-temo' and 'ii-desuka',

respectively, 2) considering the semantic features of the first three J-words, and taking the fourth J-word 'rokuonsuru' as V, the valency pattern:

$N[\text{loc}]'de' + N[\text{act,abs}]'wo' + V,$

can be applied to them, 3) there are two possible connections: 'kaigi-no yousu' and 'kaijyou-no kaigi', and 4) considering both 2) and 3) together, sentences:

'kaijyou-(de, no) kaigi-no yousu-wo rokuonshi-(te, temo) ii-desuka',

can finally be derived.

In a similar way, (b) shows how the following sentence is to be processed:

'genkou-wa itsu-madeni okure-ba yoroshii-desuka',

(By what time may I send the manuscript?).

Here, 1) 'okure' is combined with 'yoroshii' by the F-words 'ba' and 'desuka', to yield 'okure-ba yoroshii-desuka', 2) analyzing the semantic features of the nouns 'genkou' and 'itsu' and the presence of the verb 'okuru', the following valency pattern is applied:

$N[\text{con/-tra}]'ha' + N[\text{tim/pro}]'madeni' + V.$

3) Since the nouns 'genkou' and 'itsu' cannot be connected by the F-word 'no', 4) the following sentence:

'genkou-wa itsu-madeni okure-ba yoroshii-desuka',

is finally obtained.

### 5 An experiment to produce sentences from valency patterns

Using the valency patterns obtained from collected dialogue sentences, we carried out an experiment of producing sentences. Of the total of 256 interrogative sentences, 146 were used in determining valency patterns. The number of verbs was 65 and that of nouns was 229. In total, 669 valency patterns were prepared (10.7 patterns for each verb on the average).

In addition to the collected dialogue sentences, we prepared 70 test questions. For these interrogative sentences, we carried out a sentence-producing experiment. The results of this experiment are shown in

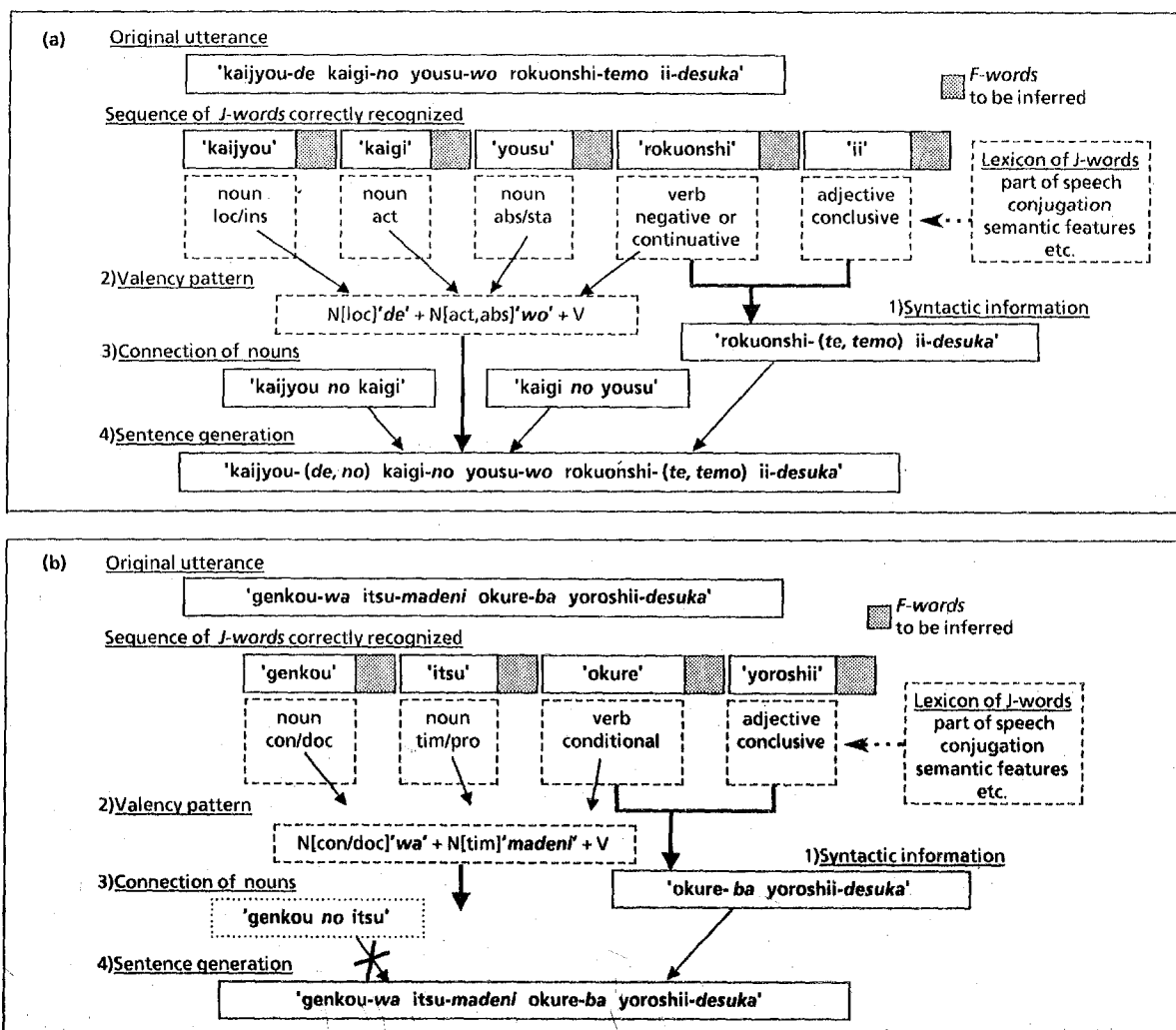


Figure 2 Process of inferring F-words.

Table 2. In this experiment, the sequence of J-words of each test sentence was input, and a complete sentence including F-words was output. The correct answer rate in Table 2 is the percentage of all the output sentences that were consistent with the input sentences. At the first trial, 64.3% correct sentences were produced from the prepared valency patterns.

Table 2 Sentences obtained using valency patterns

	Trial	1	2	3	4	5	6
Number of candidate sentences	1	12	14	15	16	16	16
	2	14	20	25	24	24	24
	3	0	0	0	5	5	5
	4	3	3	3	3	3	3
	5	7	7	7	7	7	7
	6	6	6	6	6	7	5
	7	0	0	0	0	0	4
	8	2	2	2	2	2	3
	9	0	0	0	0	0	0
	10	1	1	1	1	1	1
Correct sentences (total)		45	53	59	66	67	70
Correct answer rate (%)		64.3	75.7	84.3	94.3	95.7	100
Number of candidate sentences	0	8	0	0	0	0	0
	1	6	6	0	0	0	0
	2	6	6	7	0	0	0
	3	1	1	0	0	0	0
	4	0	0	0	0	0	0
	5	1	1	1	1	0	0
6	3	3	3	3	3	0	
Incorrect sentences (total)		25	17	11	4	3	0
Valency patterns (total)		669	677	683	690	691	694

The upper half of Table 2 shows the number of candidates in the trials where some of the output sentences were correct. For example, the number 7 shown for the first trial in the line corresponding to 5 candidate sentences means the number of candidate sentences was 5, and that 7 of the 70 test sentences were correct ones.

The lower half of Table 2 shows the number of test sentences in trials where no candidate sentence was produced or where none of the candidate sentences produced were correct.

The figures for the second and subsequent trials in Table 2 show the change in the correct answer rate when additional valency patterns were used to increase the incidence of correct sentences. In this experiment, enough valency patterns were added so that the sixth trial always produced correct sentences.

All of the test sentences used in the above experiment were simple interrogative sentences. As shown in Table 2, the use of valency patterns allows easy

production of a complete sentence from a given sequence of J-words.

## 6 Inferring an omitted verb

The verb in a given sequence of J-words has an important role in this method because it allows the selection of a correct valency pattern. It would be difficult to proceed by this method if the verb is omitted for some reason, such as speech recognition failure, or if it were originally omitted as is often the case in Japanese dialogue.

However, in this restricted domain, nouns with particular semantic features are often related to particular verbs. For example, as shown in Figure 3, in sentences which contain a noun with the semantic feature of concrete/document, the noun + F-word 'wo' tends to be accompanied by the verb 'okuru'('send'), 'kaku'('write') or 'motsu'('have'), and the noun + F-word 'ni' tends to be accompanied by the verb 'kinyuusuru'('enter').

This suggests the possibility of inferring an omitted verb from the nouns by inversely applying a suitable valency pattern. In fact, the definition of a valency pattern can be generalized as follows:

$N[\text{sem } 1g/\text{sem } 1s] + N[\text{sem } 2g/\text{sem } 2s] + \dots + V[\text{v-class}]$ , where  $V[\text{v-class}]$  denotes a verb belonging to verb class 'v-class'. This valency pattern can be used to infer a verb  $V[\text{v-class}]$  associated with a set of nouns  $N[\text{sem } 1g/\text{sem } 1s]$ ,  $N[\text{sem } 2g/\text{sem } 2s]$ , etc..

This is schematically illustrated in Figure 4. When there is a verb group A (consisting of verbs which are inferred when a certain F-word is added to noun A) and a verb group B (consisting of verbs which are inferred when a certain F-word is added to noun B), the common area of these two groups indicates the verbs which are inferred from the valency pattern containing noun A and noun B.

For example, in a sentence which contains a noun

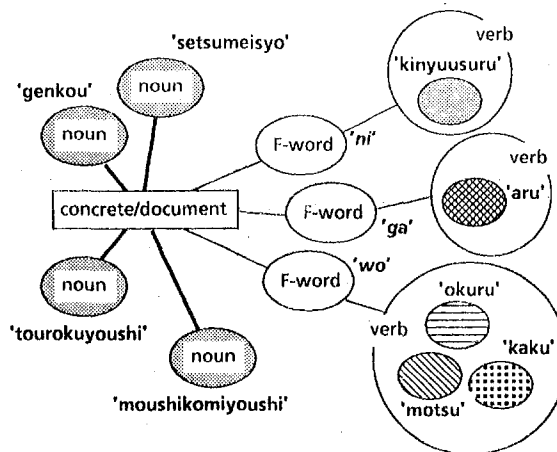


Figure 3 Frequent combinations of nouns and verbs

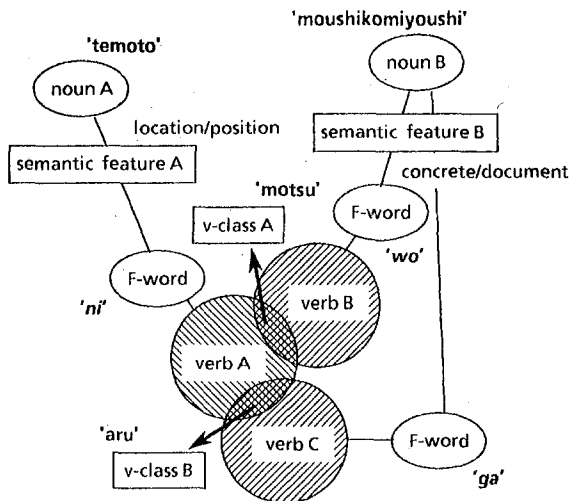


Figure 4 Verbs defined by multiple nouns

with the semantic features of concrete/document and a noun with the semantic features of location/position, verbs such as 'aru'('be') and 'motsu'('have') tend to be selected, and F-words specific to these verbs are chosen.

Table 3 shows the number of verbs which are inferred from a given sequence of nouns using the valency patterns described in Chapter 5. Since valency patterns were prepared for 65 verbs, the verbs are inferred from these 65.

The columns in Table 3 show the number of verbs inferred. The lines in the same table show the number of nouns in the valency patterns. For example, when the number of inferred verbs is 5, there are 6 valency patterns where 5 verbs are inferred; and of these 6 patterns 1 has one noun, 4 have 2 nouns, 1 has 3 nouns and none have 4 nouns.

In counting the number of inferred verbs, only verbs having a valency pattern consistent with given valency patterns are counted. When the noun of a valency pattern bears a specification as to the upper-level general semantic features but no specification as to the lower-task-dependent semantic features, the verbs of the valency patterns bearing a specification as to the lower-level semantic features are counted. Conversely, for valency patterns where the lower-level semantic feature is specified, the verbs bearing no specification as to the lower-level semantic features are not counted. For example, in the following two valency patterns, the verbs V1 and V2 are inferred from the pattern (a), while only the verb V2 is inferred from the pattern (b).

- N[con] + N[loc] + V1 (a)  
 N[con/doc] + N[loc] + V2 (b)

As can be seen in Table 3, only one verb was inferred in more than 50% of the valency patterns. In 90% of the remaining valency patterns where multiple verbs were inferred, the number of verbs inferred was 6 or less. These

Table 3 Verbs inferred using valency patterns

A \ B	1	2	3	4	Total
1	6	77	29	5	117
2	7	17	3	0	27
3	2	13	1	0	16
4	3	4	0	0	7
5	1	4	1	0	6
6	0	7	0	0	7
7	1	4	0	0	5
8	0	2	0	0	2
9	0	0	0	0	0
10	2	1	0	0	3
11	2	0	0	0	2
12	2	1	0	0	3
16	1	0	0	0	1
18	1	0	0	0	1
21	1	0	0	0	1
29	1	0	0	0	1
Total	30	130	34	5	199

A: Number of verbs inferred

B: Number of nouns in valency patterns

results indicate that, in a restricted task domain, the semantic features of the preceding nouns and valency patterns allow a fairly restricted number of candidate verbs to be inferred.

## 7 Conclusions

As a first step toward a better interface between speech recognition and machine translation, a method which is particularly useful for Japanese sentences was proposed to infer F-words for a given sequence of J-words.

In a restricted task domain, the most appropriate F-word can be inferred from a given sequence of J-words if the task-dependent semantic features of nouns are preset and the information of valency patterns is utilized.

In addition, the results of this study suggest that correct verbs can be inferred from valency patterns.

The authors are now evaluating the effectiveness of the procedures proposed in this paper by applying them to actual results of speech recognition.

## Acknowledgment

The authors are deeply grateful to Dr. Kurematsu, the president of ATR Interpreting Telephony Research Laboratories, and all the members of ATR Interpreting Telephony Research Laboratories for their constant help and encouragement.

## References

- [1] Hayes, P.J. et al., "Parsing Spoken Language: A Semantic Caseframe Approach", COLING 86.
- [2] Niedermair, G., "Divided and Valency-Oriented Parsing in Speech Understanding", COLING 86.