Speech Recognition using Valency Patterns

Jin'ichi Murakami, Ryo Saeki, and Satoru Ikehara

Dept. of Information and Knowledge Eng., Tottori University,
4-101 Koyama-Minami, Tottori 680-8550, Japan
{murakami| saeki | ikehara}@ike.tottori-u.ac.jp

abstract

We examined the effectiveness of the valency patterns when they were included in a Japanese sentence speech recognition algorithm. These valency patterns contained detailed information on case elements and structural usages of 6,000 Japanese predicates. The patterns were extracted from Nihongo Goi Taikei published by Iwanami [4].

Nihongo Goi Taikei is a large dictionary containing a total of 300 thousand entries (Japanese words) and about 14,000 valency patterns and 2,710 types (groups) of noun.

We used a speech recognition system that employed a bigram model to generate the N-best (8 best) output sentences. Next, we selected candidate sentences by using the valency patterns. Finally, we evaluated the changes in the sentence recognition rate when the valency patterns were used.

The results of experiment were that the rate improved for 23 of 70 sentences and the rate of the first candidate was 67%. This means that the valency patterns are useful for Japanese speech recognition.

Keywords: speech recognition, sentence recognition rate, the valency patterns, N-gram

1 Introduction

A language model is primarily based on two types of information. One is grammatical and syntactic information, and the other is stochastic and statistical information. Currently, a word N-gram model has usually been used in speech recognition system[1]. The word N-gram model is local statistical information on language and does not include syntactic information.

Conventionally, some language models of speech recognition systems use context-free grammar [2] or network grammar [3]. However, these language models have not been proved whether these models are effective or not. One of the reasons is that the number of rules and vocabulary used by them are limited. In this paper, we will examine the effectiveness of a large number of the valency patterns in Japanese sentence speech recognition.

We used the valency patterns of Nihongo Goi Taikei, a dictionary published by Iwanami [4]. These valency patterns contain detailed information on case elements and structural usages of 6,000 Japanese predicates. The dictionary contains a total of 300-thousand Japanese words, about 14,000 valency patterns, and 2,710 semantic categories of noun that are composed in a 12-layer hierarchy.

First, we used the speech recognition system that employed a bigram-based language model to generate the N-best (8 best) output sentences. Next, we selected candidate sentences using the valency patterns. Finally, we evaluated the change in the sentence recognition rate when the valency patterns were used.

Note on the paper; the strings enclosed by 「 and 」 are the Japanese Kanji-Kana expression, and the strings enclosed by " and " indicate the Japanese pronunciation for the sentence. The strings enclosed by ' and ' are the English translation. UGS means an ungrammatical sentence or a semantically incorrect sentence.

2 Speech Recognition using Valency Patterns

The valency patterns [4] were from a valency dictionary and a word dictionary. The dictionaries are described in detail below.

2.1 Valency Dictionary

The valency dictionary is a collection of sentence structures for 6,000 Japanese verbs and adjectives in the form of valency patterns. Let's look at the Japanese word 膨れる (Fukureru). This word means "to swell or expand" in English.

膨れる (*Fukureru*)

1. $N1$ が $N2$ に 膨れる (*"Fukureru"*)
   (' $N1$ swell to $N2$ ')

   $N1$（134 客 'Guest' 114 仲間 'Associate' 121 相手 'Partner' 533 具体物 'Concrete' 2585 数量 'Quantity' ）
   $N2$（2585 数量 'Quantity' ）

2. $N1$ が $N2$ で 膨れる (*"Fukureru"*)
   ('$N1$ bulge with $N2$')

   $N1$（821 衣服 'Clothes' 908 袋 'Bag' 820 衣服 'Clothes'）
   $N2$（533 具体物 'Concrete' 1190 金銭 'Money' ）

⋮ ⋮

The first line of each entry gives the Japanese valency pattern. Arguments (case elements) are indicated by $N1$, $N2$, etc.

The second line indicates the selective restrictions for the arguments. That is, it specifies nouns that may be instantiated by $N1$ and $N2$ according to their semantic category. Sometimes, the semantic category is complex.

This example shows that "134 客 (= guest)", "114 仲間 (= associate/member)", ... , "2582 列 (= line/queue)" can be used for $N1$, and category "2585 数量 (= quantity)" for $N2$.

2.2 Word Dictionary

The word dictionary contains 2,710 semantic categories and composed of a 12-layer hierarchy. It defines 300,000 Japanese nouns.

For example, let's look at the entry for the word "議員 (councilor)".

議員 (*"giiN"*) 名詞 'Noun'

119 成員 'Member' 260 政治家 'Politician'

The first line shows the Japanese Kanji characters for the word "議員", its pronunciation, and its part of speech. The next line indicates the semantic categories for this word. When the word belongs to several categories, these are listed in order from the most fundamental to more derivative. In this example, the categories are "119 成員 ('Member')" for the fundamental and "260 政治家 ('Politician')" for the derivative. The 12-laye hierarchy of this word dictionary is figured in Fig. 1.

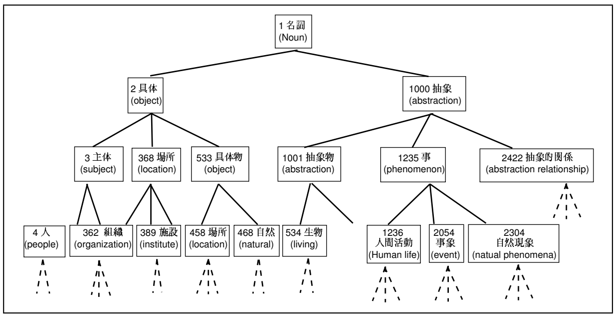

Figure 1: Word dictionary

2.3 How to Use Valency Patterns

Table 1 shows an example Japanese sentence of the speech input data.

Table 1: Input Sentence (Example)

「私は彼の高校を受験する.」
"Watashiwa kareno koukouwo jyukeNsuru."
' I will take an entrance examination for his high school.'

We used the word bigram model to generate the 2-best output sentences (table 2).

Table 2: Examples of Output Sentences

order	output sentence
1	「私からの高校と受験する.」 (UGS) *"watashikarano koukouto jyukeN suru." (UGS)* 'I take examination with a senior high school from me.' (UGS)
2	「私は彼の高校を受験する.」 *"watashiha kareno koukouwo jyukeN suru."* 'I take an examination to his senior high school.'

Table 3 shows the valency patterns for the Japanese verb「受験する」('take an examination'). In this table, Japanese or English words in parentheses mean the semantic categories for the N's.

Table 3: The Valency Patterns for 「受験する」('take an examination')

1	「 $N1$(人) [は/が] $N2$(362 組織) を 受験する」 ' $N1$(human) [ha/ga] $N2$(362 Organization) take an examination '
2	「 $N1$(人) [は/が] $N2$(1426 試験) を 受験する」 ' $N1$(human) [ha/ga] $N2$(1426 Examination) take an examination '
3	「$N1$ (人) [は/が] 受験する」 ' $N1$ (human) [ha/ga]take an examination '

Because none of the valency patterns in Table 3 match the first candidate, it must be discarded as an ungrammatical sentence. Whereas, since one of the valency patterns in Table 3 matches the second candidate, it becomes the first candidate.

3 Experiment using Valency Patterns

3.1 Procedure

The procedure of the experiment was as follows.

1. Input sentence

   Simple sentences were recorded as speech input data.

2. Recognize speech data using word bigram model

   The 8-best candidates were generated using the speech recognition system with the word bigram model.

3. Use the valency patterns to select candidates

   The valency patterns were used to analyze the 8 best candidates for further selection.

4. Evaluate the sentence recognition rate.

   The sentence recognition rate was evaluated.

The following sections details each of the above steps.

3.2 Input Sentence

The speech data were simple sentences included in IPAL [7]. They were standard and very popular Japanese single sentences. We extracted 70 sentences from IPAL for the experiment and one male speaker read them aloud for the recording of the experiment. Table 4 shows examples of input sentences.

The valency patterns do not accept all possible Japanese sentences. Thus, we had to delete such input sentences and out of problems, so that all the input sentences would be accepted by the valency patterns.

Table 4: Input Sentences (Samples)

1	「彼は試験に備えている.」 *"karewa shiken'ni sonaeteiru."* ' He is preparing for his examination.'
2	「私は水で手を洗った.」 *"watashiha mizude tewo aratta."* ' I washed my hand in the water.'
3	「私は子供に名前をつけた.」 *"watashiha kodomoni namaewo tsuketa."* 'I have named my child.'
4	「私はその絵に心を奪われた.」 *"watashiha sonoeni kokorowo ubawareta."* ' I lost my heart to that picture.'

3.3 Speech Recognition System

We used a tree-trellis type speech recognition system, which was developed in our laboratory. The Hidden Markov Kits [5] was used to learn phone HMMs. Speaker-independent HMM models were prepared using the ATR A-set word database (containing 5240 words for a set) and speech data of 10-males sets. Phone models were also adapted using the ATR B-set (503 sentences uttered by the same speaker). The word bigram model was developed with the Mainichi newspaper articles published during 1993 (about 100,0000 sentences) and 70 speech input sentences.

3.4 Evaluations

We evaluated the effectiveness of the valency patterns by comparing two sentence recognition rates.

1. The sentence recognition rate using word bigram models only.
2. The sentence recognition rate using both the valency patterns and word bigram models.

4 Results of Experiments

4.1 Example of Output Sentences (Success)

Table 5 shows a successful example of an input speech sentence. Table 6 shows the output sentences and table 7 shows the the valency patterns of the verb 「くれる」('give').

Table 5: Input Sentence (Success)

「母が私に新しい本をくれた」
"Haha ga watashini atarashii hoN wo kureta "
'My mother gave me a new book.'

The first candidate was eliminated because none the valency patterns for the verb 「くれる」('give') could be applied. The second candidate was eliminated because none of the valency patterns for the 「遅れる」('to be late') could be applied. The third and fourth candidates matched one of the valency patterns for the verb 「くれる」('give').

To summarize, based on the valency patterns, the first and second candidates were eliminated and the fourth candidate was placed in the second position.

Table 6: Example of Output Sentence (Success)

order	output sentence
1	歯が私に新しいのんをくれた (UGS) *"Ha ga watashi ni atarashii noN wo kureta (UGS) "* 'The tooth gave me new non. (UGS) '
2	歯が私のに新しいほど遅れた (UGS) *" Ha ga watashinoni atarashiihodo okureta (UGS)"* 'It was late, so that the tooth was new to mine. (UGS) '
3	母が私に新しい方をくれた *" Hahaga watashini atarashii katawo kureta."* 'The mother gave me the newer one.'
4	母が私に新しい本をくれた *"Hahaga watashini atarashii hoNwo kureta."* 'The mother gave me a new book.'

Table 7: The Valency Patterns for 「くれる」(give, strike, etc ')

1	「 $N1$ (3 主体) が/から $N2$ (388 場所 533 具体物 1001 抽象物) を $N3$ (3 主体) に くれる 」 ' $N1$ (3 object) give $N2$ (533 locate, 550 object, 1001) to N3 (3 object) '
2	「 $N1$(4 人) が $N2$(4 人) に/へ $N3$ (" ゲンコツ ") を くれる 」 $N1$(4 people) strike $N2(4people)$ with $N1's$ $N3$(the knuckles)

4.2 Example of Output Sentences (Failure)

Table 8 shows an example of failure of an input sentences, and table 9 shows the candidate (output) sentences.

Table 8: Input Sentence (Failure)

「彼は彼女に風邪をうつした」
"Karewa Kanojyoni Kazewo utushita"
' He gave his cold to her. '

Table 9: Example of output sentence (Failure)

order	output sentence
1	「彼は彼女に風をうつした」 *"Karewa Kanojyoni Kazewo utushita"* ' He gave his wind to her. '
2	「彼は彼女に風邪をうつした」 *"Karewa Kanojyoni Kazewo utushita"* ' He gave his cold to her. '

Table 10 shows the the valency patterns of the verb 「うつした」('transfer'). The first candidate matched one of the valency patterns of the verb 「うつした」('transfer'). Therefore, it was not removed. The second candidate is a correct sentence. As a result, the sentence recognition rate remains the same. Thus, this sentence was not chosen.

4.3 Sentence Recognition Rate

Our experiments showed that the sentence recognition rate improved for 23 out of 70 sentences when both word bigrams and the valency patterns were applied. Moreover, the sentence recognition rate for the first candidate increased from 50% (35 sentences, word bigrams only) to 67% (47 sentences,

Table 10: The Valency Patterns for 「うつした」 (‘transfer , give, etc ’)

1	「 $N1$ (4 人) が $N2$ (2 具体物) を $N3$(2 具体物) から/より $N4$ (場所) に/へ 移す」 ’ $N1$ (4 people) transfer $N2$ (2 object) from N3 (2 object) to N4 (2610 place) ’
2	「 $N1$(4 人) が $N2$(2416 病気) を $N3$（4 人）に 移す $N1$(4 people) give N3 (2416 disease) $N2$ (4 people)

Table 12: Example of the Sentence Ending in a Noun

Input sentence:
彼女は英語を始めた *“Kanojyo wa eigo wo Hajimeta. ”* ‘She begin to study English.’

The first candidate:
彼女は英国は 自宅. (UGS) *“Kanojyo ha eikoku wo jitaku (UGS)”* ‘She is Enlish is home.’ (UGS)

word bigrams and the valency patterns).

Table 11 shows the sentence recognitions rates for each candidate when only word bigrams were applied and both word bigrams and the valency patterns were applied. The numerators in the parentheses show the number of correct sentences and the denominators the total number of sentences investigated.

Table 11: Sentence Recognition Rate

order (candidates)	word bigram only	word bigram and valency patterns
1	50% (35/70)	67% (47/70)
∼ 2	71% (50/70)	81% (57/70)
∼ 4	84% (59/70)	93% (65/70)
∼ 8	94% (66/70)	99% (69/70)

5 Considerations

5.1 Sentence Ending in Noun

As mentioned above, our experiments show that the sentence recognition rate improves in 23 out of 70 sentences. However, we found that, in 8 of these 23 sentences, the sentence format ending in a noun contributed to the improvement of the rate, even though the sentence was grammatically, semantically incorrect (UGS). Below is an example of a sentence ending in a noun.

Such sentences can be easily eliminated by using a simple rule that selects only sentences ending with a verb or verb phase, instead of using the valency patterns.

In this study, we used the word bigram model for the speech recognition program and calculated word bigrams using the sentences collected from newspaper issues published in one year [6]. These issues contained many headlines that usually ended with a noun. Accordingly, many sentences ending with a noun were used as data for calculating word bigrams. Consequently, the text data for computing word bigrams should be such that the end of the sentence is noun.

We should remove the running titles of the newspapers before calculating the word bigram models.

5.2 Unknown Words

Upon examining the results, we found that ”unknown words” existed in the candidate sentences. The first sentence in Table 13 contains one of these words (see Table 13). The word のん” (noN) ” has almost no meaning. However, this word was accepted by the system because the valency patterns can be applied even in this case. In other words, this sentence containing this word was not interpreted as ungrammatical.

Table 13: Example of Unknown Word

母が私に新しい のん をくれた *“ Hahaga watashini atarashiinoNwo kureta. ”* ‘The mother gave me the newer None’

We think that the overall recognition rate can be further improved by using the valency patterns that can take into account these unknown words.

5.3 Lack of Valency Pattern

In our experiments reported here, all the input sentences were accepted by the valency patterns. However, the valency patterns does not always accept the sentence. In some cases, input sentence were not accepted the valency patterns. Table 14 and 15 shows an example of such a sentence.

吹く (*Fuku*)

Table 14: Input Sentence (Sample)

「風は東から吹いている.」
“kazewa higashikara fuiteiru.”
‘ The wind has blown from east.’

1. $N1$ が $N2$ から／より $N3$ に吹く (*“Fuku”*)
   (‘ $N1$ swell to $N2$ ’)

   $N1$（2373 風 ‘Wind’)
   $N2$（388 場所 ’Locate’ 2610 場 ‘Scene’)
   $N3$（388 場所 ’Locate’ 2610 場 ‘Scene’)

In this case, the valency pattern of 吹く (*Fuku*) needs $N3$. However, $N3$ does not appear in the input sentence.

Another exmaple is as follows.

Table 15: Input Sentence (Sample)

「このクラスの授業は英語で行われる.」
“kono kurasuno jugyouha eigode okonawareru.’
‘This class is conducted by English.’

行う (*Okonau*)

1. $N1$ が $N2$ を 行う (*“Okonau”*)
   (‘ $N1$ carry out $N2$ ’)

   $N1$（3 主体 ‘Subject’)
   $N2$（459 立案 ’plan’ 1826 規制 ’regulation’ 1155 制度 ’system’)

In this case, $N1$ of the valency pattern of 行う (*Okonau*) needs （3 主体 ‘Subject’) . However, $N1$ of the input sentence (授業 ’Lesson’) is not included in these semantic categories.

25 sentence out of 100 sentence did not accept the valency patterns. The coverage of the valency pattern was 75%. This rate is probably a little low. Thus we must correct the valency patterns more.

5.4 Unigram of the Valency Patterns

Our experiments used the valency patterns as a binary operator that decided whether a sentence was grammatical or not.

On the other hand, we could have used the unigram of the valency patterns. This method would have a natural affinity with the word bigram models. Also, it may improve the sentence recognition rate. Thus we will try this method in the future.

5.5 Challenge of Compound and Complex Sentences

Our experiments studied only single sentences. These are the simple usage for valency patterns. In the future, we will study compound sentences and complex sentences. It is no so easy to apply valency patterns to such sentences because they have more than one verb.

6 Conclusion

We examined how the valency patterns can contribute to the improvement of speech recognition accuracy. The valency patterns are a collection of semantic information, describing the functional relationship of noun groups to the main verb of a sentence. As the result of our experiments, the recognition rate improved for 23 of 70 sentence, and the rate for the first candidate increased from 50% to 67%. This means that it is effective to use the valency patterns for speech recognition.

In the future, we may need to research the effectiveness of the valency patterns that incorporate more complicated sentences, such as compound or complex sentences.

References

[1] Lee, Kai-Fu. (1988) “Large-Vocabulary speaker independent continuous speech recognition: the SPHINX system,” 15213 CMU-CS-88-148.

[2] T. Hanazawa, K. Kita, S. Nakamura, T. Kawabata, K. Shikano: “ATR HMM-LR Continuous Speech Recognition System”, In Readings in Speech Recognition, A. Waibel and K-F. Lee (Eds.), Morgan Kaufmann Publishers, pp. 611-614, 1990.

[3] Lowerre, B.T, “The HARPY Speech Recognition System”, PHD Thesis CMU, April, 1976.

[4] S. Ikeahera, M. Miyazaki, S. Shirai, et al., “Nihongo Goi Taikei”, Iwanami (1997).

[5] “Hidden Markov Model Toolkit”, http://htk.eng.cam.ac.uk.

[6] “Mainichi News Paper(1993.1～1993.12)“, THE MAINICHI NEWSPAPERS.

[7] “IPA Lexicon of the Japanese Language for computers (IPAL)(1987,1990,1997)“