ANALYSIS FOR STRUCTURE OF MONOLOGUE

Hideki Kashioka

ATR Spoken Language Communication Research Laboratories
2-2-2 Hikaridai "Keihannna Science City" Kyoto 619-0288 Japan

ABSTRACT

Monologue is an important knowledge source if we would be able to access to them conveniently and get interesting information from them. In order to realize convenient access to monologue archives, each monologue in those archives should be annotated various kinds of information, i.e. word boundaries, part-of-speech tag, phrase structure, dependency structure, discourse structure, and so on. In this paper, we pay attention to annotate discourse structure by using dependency analysis. We propose the monologue structure analysis method to get a kind of discourse structure. Our proposed method has two features. The first feature in our method is a processing unit that based on clause boundary unit. The sentences in monologue tend to be long, and the boundaries between sentences are not clear. Therefore if we regard whole utterance in a monologue as one sentence, the system can get the dependency structure by dependency parsing with these long sentence considering clause boundary units. The second feature is word form change at the end of sentence for concatenation with whole utterance in a monologue into one sentence.

1. INTRODUCTION

We usually sit in on our interesting lectures, watch TV commentary programs, and hear a speech out for getting knowledge in our intelligent activity. Some of these monologues are recorded into video/digital-device and we can reuse them. Recently we can access easily amount of these monologues via WEB. However we come up against selection with suitable data for our current purposes. The current these data is not familiar for us because we have little information for the detail of contents. The rich information for these monologue data is required when we construct monologue archive-database. Of course, one of the useful information is all transcribed text, and annotated transcribed text is more useful to access the suitable monologue. In this paper, we propose the analysis method with monologue for annotation.

2. PROCESSING BASIC UNIT

In this paper, we use a transcription corpus with a TV commentary program called "ASU WO YOMU" broadcasted by NHK (the Japan Broadcasting Corporation). Each program consists in a 10-minutes presentation of a current event by a commentator. Each program contains about 60-70 sentences, and each sentence has about 30 words on average[1]. Most of the sentences are complex or compound.

When we would like to analyze monologue structure, we should be considering with a basic unit with constructing for the monologue structure. In written text, we prefer a sentence as a basic unit. The sentence is clear to recognize in text because the period indicates the sentence boundary. In dialogue, an utterance is a clear unit as speaker's rotation. An utterance sometimes includes some sentences. Then, in monologue, an utterance boundary is not clear because utterance in monologue is not segmented by turn (speaker changing). Speaker would be controlled his/her utterances with an ambiguous expression at a candidate for the end of the sentence. So speaker decides using sentence ended expression or not by the reaction of audience, speaker's feeling, and an environment with his situation. Then sentence boundaries are not clear, too. Therefore monologue includes various length sentences[2], and these sentences sometimes have complicated relationships between continuous sentences. Anyway, it is necessary to define some short and suitable units instead of whole sentences for processing monologues.

When we stand the view point with grammatical side, there are several units as word, phrase, clause, sentence, etc. A sentence is one of an important unit for natural language processing. Clause is also an important unit like sentence and it is almost included in at least two elements (i.e. subject and verb). The clause boundaries can be detected with a POS sequence for a few words sequence. Thus, this paper proposes the analysis method by using dependency parsing based on clause boundaries unit.

[1]The average length of a sentence in a monologue is much longer than that in a dialogue.

[2]In monologue, sentence boundaries are not clear. However transcription includes in the sentence boundaries that were judged by human.

3. STRUCTURE OF MONOLOGUE

The purpose of discourse structure analysis would be figured out the relationships among sentences [1][2]. In this section, I describe the monologue structure that is output with our proposed method. The basic idea is very simple; the monologue structure is dependency structure like sentence analysis with dependency parsing. Current dependency parser would be able to analysis complex sentence. For example, following Japanese sentence (1) is a complex sentence.

(1) SENGETSU (last month) ATARASII (new)
GEEMU (game) GA (*particle*)
HATSUBAI (sold) SARE (*aux-verb*)
OOKU (many) NO (*particle*)
WAKAMONO (young) GA (*particle*)
KOUNYUU (buy) SITEIMASU (*aux-verb*).
"Last month, new game was started to sold, and many young people bought it."

The dependency parsing result of this sentence is shown in figure 1. This result shows the clause "SENGETSU ... HATSUBAI SARE" was modified "KOUNYUU SHITEIMASU".

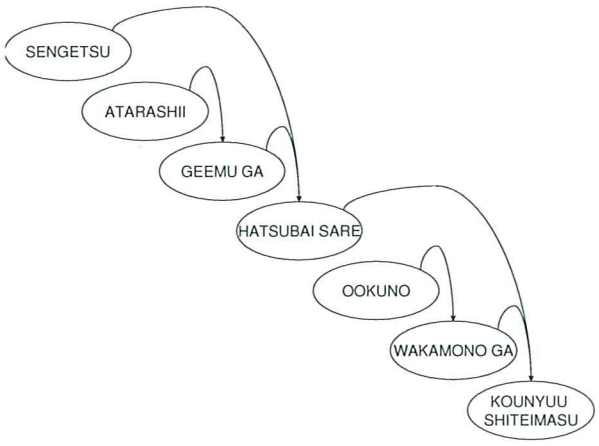

Fig. 1. Analysis result of complex sentence

On the other hand, this sentence could be divided into following two sentences (2-1, 2-2); the meaning of those sentences is same as one sentence (1).

(2-1) SENGETSU (last month) ATARASII (new)
GEEMU (game) GA (*particle*)
HATSUBAI (sold) SARE MASHITA (*aux-verb*)
"Last month, new game was started to sold."

(2-2) OOKU (many) NO (*particle*)
WAKAMONO (young) GA (*particle*)
KOUNYUU (buy) SITEIMASU (*aux-verb*).
"Many young people bought (that new game)."

We can find two modifications in previous example sentences, (1) and (2-1, 2-2). One modification is that one sentence (1) segmented into two sentence (2-1, 2-2), and another modification is two sentences (2-1, 2-2) concatenated one sentence (1).

The parsing result of these two sentences (2-1, 2-2) does not indicate the relationship with these two sentences. Current dependency parser would be processed sentence by sentence, and did not analysis the relationship between the sentences. However, the difference between these two expressions ((1) and (2-1, 2-2)) is only the additional expressions "(SARE) MASHITA". The relationship among separated two sentences is clearly for looking parsing result with concatenated sentence. Then, if we have two sentences like (2-1, 2-2) in monologue and we would be able to concatenate these sentences into one sentence like (1), we would be able to analyze the relationships among sentences with parsing.

In Japanese if we concatenate whole sentences in the monologue into one sentence by modification of the expression that appeared at the end of sentence like this example, we can get the monologue structure as a result of dependency parsing with a concatenated sentence.

Then, we show one more example that is constructed with five sentences.

1. WATASHI(I) WA(*particle*)
   KAZE(cold) WO(*particle*)
   HIITE(catch) INAI(not)
   I did not catch a cold.

2. DAKARA(then)
   GAKKOU(school) NI(*particle*)
   IKIMASHITA(went(go))
   I went to school.

3. SHIKASHI(however)
   KAZE(cold) DE(*particle*)
   TOMODACHI(friend) GA(*particle*)
   TAKUSAN(many)
   YASUNDE(absent) IMASHITA(were(be))
   However, many friends were absent with a cold.

4. SOREDE(therefore)
   GAKKYUU(class) HEISA(closed) NI(*particle*)
   NARIMASHITA(become)
   Therefore, the class was closed.

5. SONOTAME(then)
   GOGO(afternoon) KARA(from)
   IE(home) NI(*particle*)
   KAERI(come back) MASHITA
   Then I came back home from afternoon.

These five sentences are concatenated into following one sentence that includes five clauses.

1. WATASHI WA
   KAZE WO
   HIITE INAI NODE
   I did not catch a cold.

2. (DAKARA)
   GAKKOU NI
   IKIMASHITA GA
   I went to school.

3. (SHIKASHI)
   KAZE DE
   TOMODACHI GA
   TAKUSAN
   YASUNDE ITE
   However, many friends were absent with a cold.

4. (SOREDE)
   GAKKYUU HEISA NI
   NARI
   Therefore, the class was closed.

5. (SONOTAME)
   GOGO KARA
   IE NI
   KAERI MASHITA
   Then I came back home from afternoon.

This sentence analysis result is figured in figure 2.

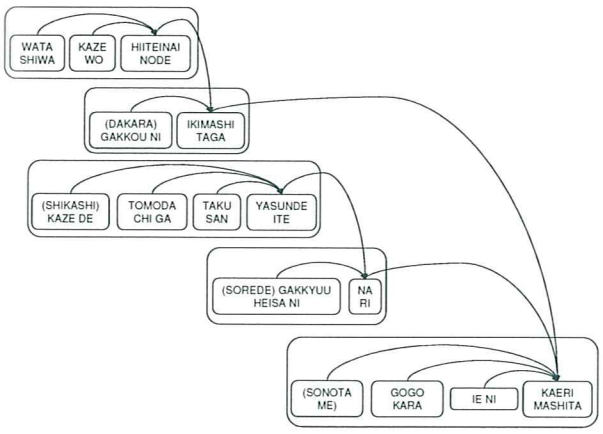

Fig. 2. Analysis result of concatenated sentence from 5 sentence

The contents of utterance with one sentence are almost same as constructed with two or more sentences.

4. PROCESSING METHOD

In previous section, we explain the idea of monologue structure by dependency parsing. Our proposed method should be based on the processing unit as clause. Because the relationship between sentences would be indicated the relationship between the last bunsetsu in the sentence and the other bunsetsu that included following sentence and these relationship seems as the relationship between the last bunsetsu in the clause and the other bunsetsu that included following clauses inside the sentence.

In order to apply the dependency parsing to long concatenated sentence, we mainly deal with two issues, one is how to concatenate sentences, and the other is how to apply parser for long sentence based on clause. In this section, we explain how to parse with long sentence and how to concatenate each sentence.

4.1. Proposed parsing method

Usually, long sentence is difficult to parse because long sentence tend to be complicated and grammatically uncompleted. Our proposed method should be treated very long sentence [3] for parsing. A lot of researcher would try to segment long sentences into several sentences, and processing with each segmented short sentences. Almost of them are not treated for relationships with segmented sentences. However, Ohno [3] proposed efficient method for parsing with long sentence. His proposed parsing method has two level parsing, one is inside-clause-level parsing and the other is outside-clause-level parsing.

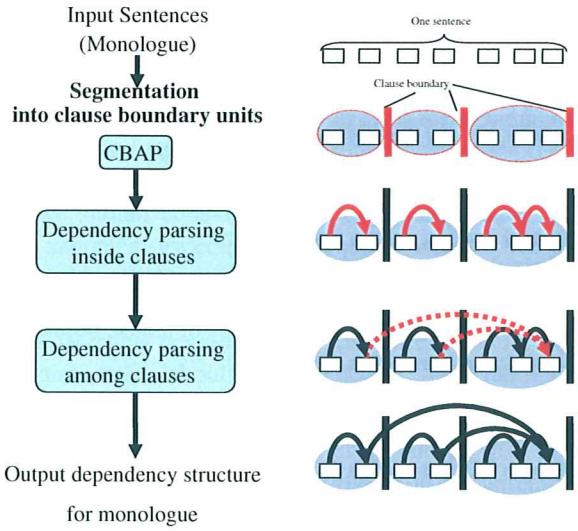

Fig. 3. the flow of two step parsing

Figure 3 shows the flow of two level parsing method. Initially an input sentence would be segmented into clause boundary unit by CBAP [4]. CBAP is the program which detects and labels every clause boundary in a Japanese sentence. This

[3] In our experiment, we use the transcriptions of TV commentary program and there are 60 sentences par program in average.

Table 1. Top 10 expression at the end of sentence("ASU WO YOMU" data)

	frequency	expression at EOS	word form change
1	865	TO * OMOI MASU	TO * OMOI MASU GA
2	444	SHI TE I MASU	SHI TE I TE
3	219	NA TTE I MASU	NA TTE I TE
4	218	RE TE I MASU	RE TE I TE
5	198	IU * KOTO DESU	IU * KOTO DE
6	183	TE KI MASHI TA	TE KI MASHI TE
7	154	SHI TE ORI MASU	SHI TE ORI
8	153	NAI DESYO U KA	——
9	150	NO DESYO U KA	——
10	134	MASE NN DESHI TA	MASE NN DE

program developed by Maruyama [5]. The core of CBAP is a set of "clause boundary annotation rules," described manually. A total of 332 annotation rules are included in CBAP. Japanese subordinate clauses can be classified by their relational meanings between clauses and by their morphological features. In CBAP, they prepared 147 types of boundary labels to represent the types of clause boundaries. The labels include a few other constituents that do not strictly represent clause boundaries but can be regarded as being syntactically independent elements, like topically elements, conjunctive, interjections, and so on. These boundary labels are useful for the relationship between clauses in our proposed method.

Next, Ohno's method proceeds the following two level parsing:

Inside-clause-level parsing
The dependency relations of a clause boundary unit inside are identified for every clause boundary unit in a sentence.

Outside-clause-level parsing
The dependency relations of which the dependent bunsetsus are the last bunsetsus of clause boundary units in a sentence are identified.

Ohno's parsing method uses the criteria for judgment of the pair of bunsetsus that has dependency relation by calculating for probability provided from the corpus. However his method would be work with other criteria, such as SVM learning, decision tree, and so on.

His method has following features:

- In each step, the candidates for dependency pairs would decrease. So totally process of parsing would be faster than conventional method.
- Incremental processing would be possible.
- Clause type information would be useful for extended analysis.

Thus, we modified his method for monologue structure analysis. Figure 4 shows the flow of our proposal analysis.

The main flow is two differences with Ohno's method. One is the module of "word form change" [4] is inserted after outside-level-parsing modules. In our proposed method, this outside-level-parsing is not really outside-level-parsing. Our method's input is multiple sentences. Then if outside-level means really outside of clause, the bunsetsu in the following sentence would include in the candidate for the modified bunsetsu. Assumedly, the relationship between the dependent pair would be closed in sentence. Therefore, our proposed method has the additional third parsing step after "word form change" module.

Ohno's method treats only one sentence, so the bunsetsu which included in candidates for the dependency pairs is absolutely in the sentence. However, our proposed method treats multiple sentences, and we have to treat the relationship between the last bunsetsu in the sentence and other bunsetsu that was included in the following sentences. We would separate these analyses from second parsing step as same reasons for second parsing step separated from first parsing steps.

4.2. Word form change at sentence ends

We show some examples of word form change at the end of sentence in section 3, those word form change are

1) "SARE MASHITA(*aux verb*)"
→ "SARE",

2) "INAI. (DAKARA)"
→ "INAI NODE (*particle conjunction*)",

3) "IKIMASHITA"
→ "IKIMASHITA GA(*particle conjunction*)"

and so on.

[4]The details of this inserted modules is explained in next section 4.2.

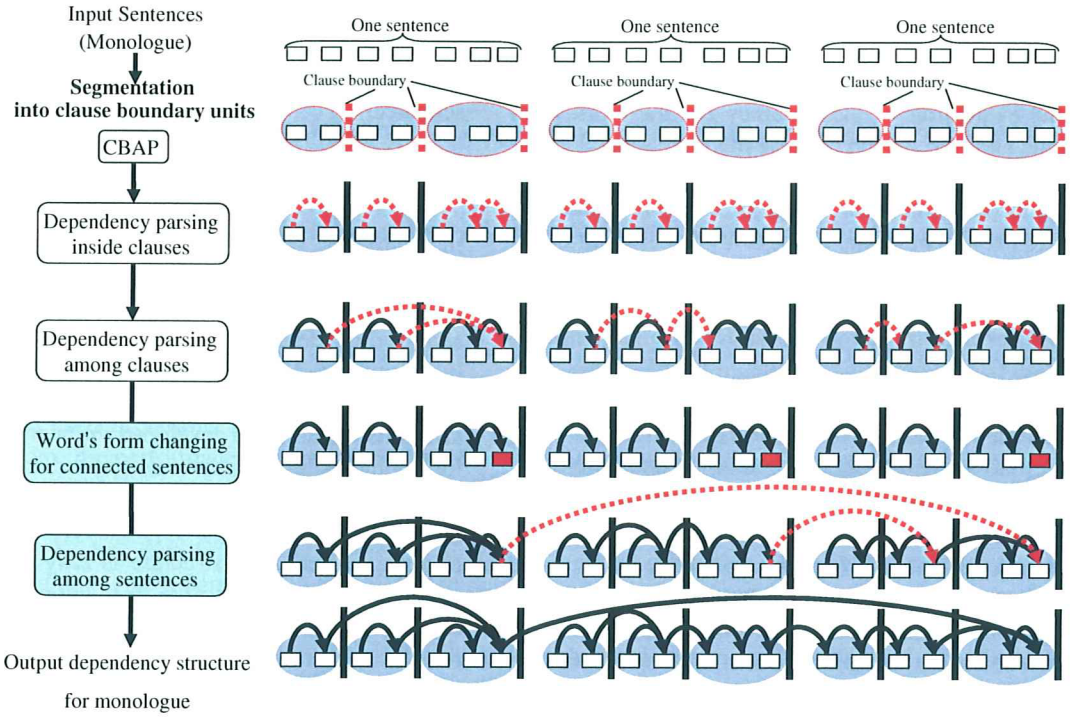

Fig. 4. the flow of our proposal analysis

Almost expression at the end of sentence would be able to connected to the next sentence with adding *particle conjunction* "GA, NODE, etc". In case of 1) and 3), these change form of the word at the end of sentence seems generally. Especially the expression "GA*particle conjunction*" is very simple expression and many people frequently used for continuing their uttered sentences.

In case of 2), the form change is depended on the following expression "DAKARA" at the next sentence. Looking over these form change, currently we try to make the rule set for word form change like CBAP rules.

In order to make rule set, we collect the expression at the end of sentence. Table 1 shows the top 10 expression constructed four-words at the end of sentence in "ASU WO YOMU" 250 program data. "ASU WO YOMU" data is a transcribed text for a TV commentary program that titled "ASU WO YOMU" of the Japan Broadcasting Corporation (NHK). Each program comments on some current social issue for 10 minutes by a different commentator.

An original expression at the 8th and 9th line in the table 1 is question form. Like these question form, it is difficult to concatenate into the following sentence by word form change. Current our answer to treat this case is inserted in the expression "TO IUKOTO DE". This expression works that the previous sentence becomes an embedded sentence. However, we should take care of parsing knowledge would be able to treat an embedded sentence.

Four-word expressions at the end of sentence are appeared about six thousand variations in 250 programs. More than 10 times appeared expressions are about two hundred variations and these two hundred variations covers about 45% of sentence ends. We wrote word form change expression to more than 10 times appeared, and then, we create the general rules from these list. In addition to make the general rules, we should consider the following expressions at the end of sentence, especially conjunction words. For example, if "DAKARA" is appeared following sentence, the end of sentence would be changed suitable conjugation and adding "NODE (*particle conjunction*)". Therefore, we currently has five general rules, about ten conjunction words modification rules, and about a hundred example rules like in table 1

These rules would not be count correctly because some of example rules would be worked as same as effect in general rules and conjunction word modification rules.

5. EXPERIMENT

We apply our proposed analysis method into "ASU WO YOMU" data. I mentioned "ASU WO YOMU" data a little in the previous section. Each program includes about 60 sentences in average. Each sentence constructs with 3 or 4 clause boundary units and sentence length is about 50 words.

The coverage of word form change at the end of sentence is almost 90% of the sentence end in "ASU WO YOMU" data.

Current example rules covered about 45% in fully "ASU WO YOMU" data. However, this coverage would be achieved higher with extended example rules. We cannot find a good way for the evaluation of the accuracy for the word form change at the end of sentence because there are several ambiguity to represent for suitable expression to concatenate two sentences into one sentences. In addition to this evaluation, the evaluation result would be changed by the evaluator who is considering with how many words in previous sentences and following sentences. Although we did not find a critical error in local expression, however, we found a little strange concatenation of sentences in our rules. One of strange concatenation is made with question and declarative sentence. In previous section, we proposed the insertion with specific expression into after the question. Then, we expect the question would become embedded sentence. In some case speaker changed the topics after the question, so the concatenation of these sentences makes strange.

Although our proposed method should treat about three thousand words in a sentence, each parsing steps are treated only short words' (or bunsetsus') sequences. In third level parser, the knowledge for decided the relationship between bunsetsus' pair is almost same as second level parser such as between the bunsetsu at the end of clause and the bunsetsu in following clauses. Therefore, Currently third level parser hired the knowledge of second level parser.

6. DISCUSSION

Our proposed method outputs the dependency structure for whole structure of monologue. Our structure as shown in figure 2 has a dependency structure and basically parser did not output the label of the relationship between bunsetsus. However, CBAP assigned the label of clause type, and clause at the end of sentence after word form change could be assigned the label of clause type that is not "end of sentence". This assigned label is indicated the label of relationship between sentences.

And dependency structure would be shown some segmentation. In section 3, I show the example with concatenation with five sentences into one sentence, and the parsing result was shown in figure 2. This result shows first and second sentences make one segments, third and forth sentences make the other segments, and these two segments are modified last sentence. Therefore, we can segment these five sentences into three segments. And this segmentation was reasonable segmentation.

Current our experiments are processed with transcription by human. So there are assigned the end of sentence by human. In real data which was output from automatic speech recognition(ASR), we would be difficult to find the end of sentence. However, CSJ [6] project[5] did detect the clause boundary unit from ASR and they assigned the candidates for the end of sentence [5]. Therefore, our method could be used these candidate for the end of sentence seems at the end of sentence.

[5]CSJ means "the corpus of spontaneous Japanese"

7. CONCLUSION

In this paper, we proposed monologue analysis method by using dependency parsing based on clause boundary units. This method has three level parsing steps, first level parser analyzed inside-clause dependency, second level parser analyzed outside-clause dependency in each sentence, and third level parser analyzed between sentence dependencies. To realize this process, we developed the word form change rules for paraphrasing the expression at the end of sentence. Our proposed method is almost same for basic technique as conventional dependency parsing. Therefore anaphora resolution by using dependency structure is easy extended into whole monologue. We would like to maintain for parsing knowledge for outside-sentence-level's dependency relation, currently we substitute outside-clause level's knowledge for outside sentence-level's knowledge. We are planning to develop the summarization system using our monologue discourse structure by this proposed method.

8. REFERENCES

[1] Bonnie Webber and Donna Byron, *Proceedings of the ACL 2004 Workshop on Discourse Annotation*, 2004.

[2] Laila Dybkjar and Wolfgang Minker, *Proceedings of the 6th SIGdial Workshop on Discourse and Dialogue*, Lisbon, Portugal, September 2005.

[3] Tomohiro Ohno, Shigeki Matsubara, Hideki Kashioka, Naoto Kato, and Yasuyoshi Inagaki, "Incremental dependency parsing of japanese spoken monologue based on clause boundaries," in *In Proc. of 9th EUROSPEECH*, 2005, pp. 3449–3452.

[4] Hideki Kashioka and Takehiko Maruyama, "Segmentation of semantic units in japanese monologues," in *In Proc. of ICSLT-O-COCOSDA 2004*, 2004.

[5] Maruyama Takehiko, Kashioka Hideki, Kumano Tadashi, and Tanaka Hideki, "Development and evaluation of japanese clause boundaries annotation program(in japanese)," *Journal of Natural Language Processing*, vol. 11, no. 3, pp. 39–68, 2004.

[6] Maekawa Kikuo, Koiso Hanae, Furui Sadaoki, and Isahara Hitoshi, "Spontaneous speech corpus of japanese," in *Proceedings of the second International Conference of Language Resource and Evaluation*, 2000, vol. 2, pp. 947–952.