Investigating the Characteristics of Causal Relations in Japanese Text

Takashi Inui and **Manabu Okumura**
Precision and Intelligence Laboratory
Tokyo Institute of Technology
4259, Nagatsuta, Midori-ku, Yokohama, 226-8503, Japan
`tinui@lr.pi.titech.ac.jp, oku@pi.titech.ac.jp`

Abstract

We investigated of the characteristics of in-text causal relations. We designed causal relation tags. With our designed tag set, three annotators annotated 750 Japanese newspaper articles. Then, using the annotated corpus, we investigated the causal relation instances from some viewpoints. Our quantitative study shows that what amount of causal relation instances are present, where these relation instances are present, and which types of linguistic expressions are used for expressing these relation instances in text.

1 Introduction

For many applications of natural language techniques such as question-answering systems and dialogue systems, acquiring knowledge about causal relations is one central issue. In recent researches, some automatic acquisition methods for causal knowledge have been proposed (Girju, 2003; Sato et al., 1999; Inui, 2004). They have used as knowledge resources a large amount of electric text documents: newspaper articles and Web documents.

To realize their knowledge acquisition methods accurately and efficiently, it is important to knowing the characteristics of presence of in-text causal relations. However, while the acquisition methods have been improved by some researches, the characteristics of presence of in-text causal relations are still unclear: we have no empirical study about what amount of causal relation instances exist in text and where in text causal relation instances tend to appear.

In this work, aiming to resolve the above issues, we create a corpus annotated with causal relation information which is useful for investigating what amount of causal relation instances are present and where these instances are present in text. Given some Japanese newspaper articles, we add our designed causal relation tags to the text segments. After creating the annotated corpus, we investigate the causal relation instances from three viewpoints: (i) cue phrase markers, (ii) part-of-speech information, and (iii) positions in sentences.

There are some pieces of previous work on analysis of in-text causal relations. However, although causal relation instances appear in several different ways, just a few forms have been treated in the previous studies: the verb phrase form with cue phrase markers such as in (1a) has been mainly treated. In contrast, we add our causal relation tags to several types of linguistic expressions with wide coverage to realize further analyses from above three points. Actually, we treat not only linguistic expressions with explicit cues such as in (1a), but also those without explicit cues, i.e. implicit, as in (1b), those formed by noun phrases as in (1c), and those formed between sentences as in (1d).

(1) a. 大雨-が   降っ-た ため 川-が   増水し-た。
heavy rain-NOM fall-PAST <u>because</u> river-NOM rise-PAST
(explicit)

b. 大雨-が   降り-、 川-が   増水し-た。
heavy rain-NOM fall-PUNC river-NOM rise-PAST
(implicit)

c. 大雨-で   川-が   増水し-た。
heavy rain-because of river-NOM rise-PAST
(noun phrase)

d. 大雨-が　　　降っ-た-。　川-が　　増水し-た。
heavy rain-NOM fall-PAST-PUNC river-NOM rise-PAST
(between sentences)

We apply new criteria for judging whether a linguistic expression includes a causal relation or not. Generally, it is hard to define rigorously the notion of causal relation. Therefore, in previous studies, there have been no standard common criteria for judging causal relations. Researchers have resorted to annotators' subjective judgements. Our criteria are represented in the form of linguistic templates which the annotators apply in making their judgements (see Section 3.2).

In Section 2, we will outline several previous research efforts on in-text causal relations. In Section 3 to Section 6, we will describe the details of the design of our causal relation tags and the annotation workflow. In Section 7, using the annotated corpus, we will then discuss the results for the investigation of characteristics of in-text causal relations.

2 Related work

Liu (2004) analyzed the differences of usages of some Japanese connectives marking causal relations. The results are useful for accounting for an appropriate connective for each context within the documents. However Liu conducted no quantitative studies.

Marcu (1997) investigated the frequency distribution of English connectives including "because" and "since" for implementation of rhetorical parsing. However, although Marcu's study was quantitative one, Marcu treated only explicit linguistic expressions with connectives. In the Timebank corpus (Pustejovsky et al., 2003), the causal relation information is included. However, the information is optional for implicit linguistic expressions.

Although both explicit expressions and implicit expressions are treated in the Penn Discourse Treebank (PDTB) corpus (Miltsakaki et al., 2004), no information on causal relations is contained in this corpus.

Altenberg (1984) investigated the frequency distribution of causal relation instances from some viewpoints such as document style and the syntactic form in English dialog data. Nishizawa (1997) also conducted a similar work using Japanese dialog data. Some parts of their viewpoints are overlapping

with ours. However, while their studies focused on dialog data, our target is text documents. In fact, Altenberg treated also English text documents. However, our focus in this work is Japanese.

3 Annotated information

3.1 Causal relation tags

We use three tags *head*, *mod*, and *causal_rel* to represent the basic causal relation information. Our annotation scheme for events is similar to that of the PropBank (Palmer et al., 2005). An event is regarded as consisting of a head element and some modifiers. The tags *head* and *mod* are used to represent an event which forms one part of the two events held in a causal relation. The tag *causal_rel* is used to represent a causal relation between two annotated events.

Figure 1 shows an example of attaching the causal relation information to the sentence (2a), in which a causal relation is held between two events indicated (2b) and (2c) . Hereafter, we denote the former (cause) part of event as e_1 and the latter (effect) part of event as e_2.

(2)　a. そして、遠方からの観光客がＧＷに入って増える。
(As the Golden Week holidays come, the number of sightseers from all over begins to increase.)

　　b. e_1 = ＧＷに入る
(The Golden Week holidays come)

　　c. e_2 = 遠方からの観光客が増える
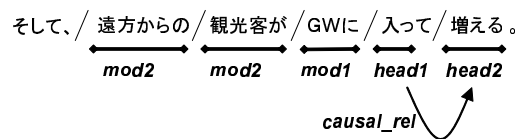(The number of sightseers from all over begins to increase)

Figure 1: An example of attaching the causal relation information

The annotation process is executed as follows. First, each sentence in the text is split to some *bunsetsu*-phrase chunks[1], as shown in Figure 1 ("/" indicates a *bunsetsu*-phrase chunk boundary). Second, for each *bunsetsu*-phrase, an annotator finds the segment which represents a head element of an event,

---

[1]The *bunsetsu*-phrase is one of the fundamental units in Japanese, which consists of a content word (noun, verb, adjective, etc.) accompanied by some function words (particles, auxiliaries, etc.).

and he/she adds the *head* tag to the segment (see also *head*$_1$ and *head*$_2$ in Figure 1). If the event has any other elements in addition to head element, the annotator also adds the *mod* tags to the segments representing modifiers to the head element (*mod*$_1$ and *mod*$_2$ in Figure 1). The elements marked with any tags which have a common suffix number are constituents of the same event: that is, the elements marked with *head*$_1$ and *mod*$_1$ tags are constituents of e_1 and the elements marked with *head*$_2$ and *mod*$_2$ are constituents of e_2. Finally, the annotator adds the *causal_rel* tag between *head*$_1$ and *head*$_2$ as link information which indicates that the corresponding two events are held in a causal relation.

When there are any cue phrase markers helpful in recognizing causal relations such as ため (because) in (1a) , the annotator also adds the *marker* tag to their segments.

3.2 Annotation criteria

To judge whether two events represented in text are held in a causal relation or not, we apply new criteria based on *linguistic test*.

The linguistic test is a method for judging whether target linguistic expressions conforms to a given set of rules. In our cases, the target expressions are two sets of *bunsetsu*-phrase chunks. Each set represents as a whole an event which can be an argument in a causal relation, such as in (2b) and (2c) . The rules are realized as linguistic templates which are linguistic expressions including several slots.

In practice, a linguistic test is usually applied using the following steps:

1. Preparing a template.

2. Embedding the target expressions in the slots of the template to form a candidate sentence.

3. If the candidate sentence is syntactically and semantically correct, the target expressions are judged to conform to the rules. If the candidate sentence is incorrect, the targets are judged non-conforming.

In this work, we prepared eighteen linguistic templates such as in Figure 2. The square brackets indicate the slots. The symbol ⟨adv⟩ is replaced by one of three adverbs しばしば (often), 大抵 (usually), or 常に (always).

$[e_1]$ （という）状態になれば、それに伴い、
⟨adv⟩ $[e_2]$ （という）状態になる。

（$[e_2]$ ⟨adv⟩ happened as a result of
the fact that $[e_1]$ happened.）

Figure 2: An example of linguistic templates

We embed two target expressions representing events in the slots of the template to form a candidate sentence. Then, if an annotator can recognize that the candidate sentence is syntactically and semantically correct, the causal relation is supposed to hold between two events. In contrast, if recognized that the candidate sentence is incorrect, this template is rejected, and the other template is tried. If all eighteen templates are rejected by the annotator, it is supposed that there is no causal relations between these two events. Note that the annotator's recognition of whether the candidate sentence is correct or incorrect, in other words, whether a causal relation is held between the two events embedded in the candidate sentence or not, is not really relevant to the author's intention.

The fundamental idea of our criteria based on linguistic test is similar to that of the criteria for annotation of implicit connectives adopted in PDTB corpus[2]. In the annotation process of the PDTB corpus, an annotator judges whether or not the explicit connective, for example, "because", relates two linguistic expressions representing events. This process is essentially the same as ours.

Three adverbs in the linguistic templates, しばしば (often), 大抵 (usually) and 常に (always), indicate a pragmatic constraint on the necessity of the relationship between any two events; the relations indicated by these words usually have a high degree of necessity. With this pragmatic constraint, we introduce an attribute to the *causal_rel* tags about the degree of necessity. For each of eighteen templates, if one judges the two target expressions as holding a causal relation by using the template with one of three adverbs, the *necessity* attribute value is added to the relation instance. If one judges the two target expressions as holding a causal relation by using the template deleting ⟨adv⟩, three adverbs, the *chance*

attribute value is added.

We assume that a target expression embedded in the slot is represented by a single sentence. If an event is represented by noun phrase (NP), the following rewriting rules are applied before embedded to the slot to transform the NP into a single sentence.

- NP → NP + する
  （ex. 停電 → 停電する）
  （ex. blackout → a blackout happens）

- NP → NP + が起こる
  （ex. 地震→地震が起こる）
  （ex. earthquake → an earthquake happens）

- NP → NP + になる
  （ex. 大雨→大雨になる）
  （ex. heavy rain → it rains heavily ）

- nominalized verb → verb
  （ex. 疲れ→疲れる）
  （ex. tiredness → someone gets tired）

If a head element of a target expression representing an event is conjugated, the head element is replaced by its base form before embedded to the slot.

3.3 Annotation ranges

Ideally, we should try to judge for tagging of the causal relation tags over all any event pairs in text. However, it seems that the more the distance between two events represented in text, the smaller the probability of holding a causal relation between them. Thus, we set a constraint on the ranges of judgements. If both two events are represented in the same sentence or two sentences adjacent to each other, we try judgements, if not, skip judgements. This constraint is applied only when tagging the *head* tag. A modifier and its head element are sometimes located in different sentences overtly in Japanese text when anaphora or ellipsis phenomenon occurs. In such cases, we add *mod* tags to the text segments anywhere in the text.

4 Data

We selected as text for annotation Mainichi Shimbun newspaper articles (Mainichi, 1995). In particular, we used only articles included on the social aspect domain. When adding the causal relation tags to the text, it is preferable that each annotator can understand the whole contents of the articles. The contents of social aspect domain articles seems to be familiar to everybody and are easier to understand than

the contents of articles included on politics, economy domain, etc.

Furthermore, in our previous examination, it is found that as the length of articles gets longer, it is getting hard to judge which *bunsetsu*-phrase chunks represent as a whole an event. This is because as described in Section 3.3, annotators sometimes need to search several sentences for modifiers of the head element in order to add *mod* tags precisely. Therefore, we focus on social aspect domain articles which consists of less than or equal to 10 sentences. After all, we extracted 750 articles (3912 sentences) for our annotation work with above conditions.

5 Annotation workflow

Three annotators have been employed. Each annotator has added tags to the same 750 document articles independently. Two annotators of the three are linguists, and the last one is the author of this paper. We denote each annotator under anonymity, *A*, *B* and *C*. After training phase for annotators, we spent approximately one month to create a corpus annotated with causal relation information. The annotation workflow is executed efficiently using an annotation interface. Using the interface, all of annotators can add tags through only simple keyboard and mouse operations. The annotation workflow is as follows.

I. *Annotation phase:* A document article is displayed to each annotator. The sentences in the document are automatically split to *bunsetsu*-phrases by preprocessing. Some kinds of words such as connectives and verbs are highlighted to draw annotators' attention to the text segments which could represent elements in causal relation instances. The annotator finds text segments which represent causal relation instances, and then he/she adds the causal relation tags to their segments as described in Section 3.

II. *Modification phase:* After each annotator finished the annotation phase for a fixed number of document articles (in this work, 30 document articles), he/she moves to a modification phase. In this phase, first, only the segments with causal relation tags are extracted from the documents such as instances in Table 1. Then,

Table 1: Examples of tagged instances

mod_1	$head_1$	mod_2	$head_2$
６階-から （sixth floor-from）	転落する （tumble）		意識不明 （lie unconscious）
川-に （river-to）	転落 （tumble）		助け上げ （help out）
二階屋根-から （roof-from）	転落 （tumble）	頭-など-を （head-ACC）	打つ （hit）
けん銃-で （handgun-with）	撃つ （shoot）	重傷-を （heavy injury-ACC）	負う （suffer）
顔-に　火傷-を （head-DAT）（burn-ACC）	負う （suffer）		重傷 （heavy injury）
重傷-を （heavy injury-ACC）	負う （suffer）		休職する （take a sabbatical leave）

the same annotator who adds tags to the extracted segments, checks their extracted causal relation instances with attention. Since the extraction is done automatically, each annotator can check all the segments to be checked. When wrong tagged instances are found, they are corrected on the moment. After checking and correcting for all the extracted instances, the annotator moves back to the annotation phase in order to annotate a new 30 document articles set.

6 Results

6.1 Total number of tagged instances

2014 instances were tagged by the annotator *A*, 1587 instances by *B*, 1048 instances by *C*. Some examples of tagged instances are shown in Table 1.

The total numbers of tagged instances of the three annotators are quite different. Although all annotators tagged under the same annotation criteria, the annotator *A* tagged to twice as many segments as the annotator *C* did. Though this difference may be caused by some factors, we assume that the difference is mainly caused by missing judgements, since the annotators added tags to a variety of linguistic expressions, especially expressions without cue phrases.

To verify the above assumption, we again asked each annotator to judge whether or not a pair of linguistic expressions representing events is holding a causal relation. In this additional work, in order to prevent the annotators from skipping judgement itself, we present beforehand to the annotators the pairs of linguistic expressions to be judged. We presented a set of 600 pairs of linguistic expressions to each of the three annotators. All of these pairs are

Table 2: Inter-annotator agreement

A	B	C	$\mathcal{S}_{mixed}$	$\mathcal{S}_n$	$\mathcal{S}_c$
1	0	0	921	632	535
0	1	0	487	487	255
0	0	1	187	134	207
1	1	0	372	230	90
1	0	1	133	92	77
0	1	1	140	107	83
1	1	1	588	270	64

the causal relation instances already tagged by one or more annotators in the main work described in the previous sections.

From the comparison between the results of the additional work and those of the main work, we found that if causal relation instances are expressed without explicit cues in text, they tend to be more frequently missed than those with explicit cues. The missing judgements on expressions without explicit cues are an important issue in the realization of more sophisticated analyses.

6.2 Inter-annotator agreement

We examined inter-annotator agreement. First, we define an agreement measure between two relation instances. Let x and y be causal relation instances tagged by two different annotators. The instance x consists of e_{1x} and e_{2x}, and y consists of e_{1y} and e_{2y}. The event e_{1x} has $head_{1x}$ as its head element. Similarly, $head_{2x}$, $head_{1y}$ and $head_{2y}$ are the head elements corresponding respectively to events e_{2x}, e_{1y} and e_{2y}. Then, we regard two instances x and y as the same instance, when $head_{1x}$ and $head_{1y}$ are located in the same *bunsetsu*-phrase and $head_{2x}$ and $head_{2y}$ are also located in the same *bunsetsu*-phrase. Using the above defined agreement measure,

we counted the number of instances tagged by the different annotators.

Table 2 shows the results. The symbol "1" in the left-hand side of Table 2 indicates that the corresponding annotator tagged to instances, and the "0" indicates not tagged. For example, the fourth row ("110") indicates that both *A* and *B* tagged to instances but *C* did not.

Let $\mathcal{S}_{mixed}$ denote a set of all tagged instances, $\mathcal{S}_n$ denote a set of all tagged instances with the *necessity* attribute value, and $\mathcal{S}_c$ denote a set of all tagged instances with the *chance* attribute value.

First, we focus on the relation instances in the set $\mathcal{S}_{mixed}$. The 1233 ($= 372 + 133 + 140 + 588$) instances are tagged by more than one annotator, and the 588 instances are tagged by all three annotators. Next, we focus on the two different contrastive sets of instances, $\mathcal{S}_n$ and $\mathcal{S}_c$. The ratio of the instances tagged by more than one annotator is small in $\mathcal{S}_c$. This becomes clear when we look at the bottom row ("111"). While the 270 instances are tagged by all three annotators in $\mathcal{S}_n$, only the 64 instances are tagged by all three annotators in $\mathcal{S}_c$.

To statistically confirm this difference, we applied the hypothesis test of the differences in population rates. The null hypothesis is that the difference of population rate is d %. As a result, the null hypothesis was rejected at 0.01 significance level when d was equal or less than 7 (p-value was equal or less than 0.00805). In general, it can be assumed that if a causal relation instance is recognized by many annotators, the instance is much reliable. Based on this assumption and the results in Table 2, reliable instances are more concentrated on the set of instances with the *necessity* attribute value than those with the *chance* attribute value.

7 Discussion

In this section, we discuss some characteristics of in-text causal relations and suggest some points for developing the knowledge acquisition methods for causal relations. Here, to guarantee the reliability of the data used for the discussion, we focus on the 699 ($= 230 + 92 + 107 + 270$) instances marked by more than one annotator with the *necessity* attribute value. We examined the following three parts: (i) cue phrase markers, (ii) the parts-of-speech of head elements, and (iii) the positions of head elements.

Table 3: The number of instances with/without cue phrase markers

with marker	219
without marker	480

Table 4: Cue phrase markers marked by annotators

marker		frequency
ため	(because)	120
で	(by)	35
結果	(result of)	5
ので	(because)	5
と	(when)	5
場合	(when)	4
ば	(if)	4
ことから	(from)	4
から	(from)	3

7.1 Cue phrase markers

While annotating the document articles with our causal relation tags, *head*, *mod*, and *causal_rel*, the annotators also marked the cue phrase markers for causal relations with the *marker* tag at the same time. We investigated a proportion of instances attached with the *marker* tag.

The result is shown in Table 3. Table 4 shows the cue phrase markers actually marked by at least one annotator [3].

It has been supposed that causal relation instances are sometimes represented with no explicit cue phrase marker. We empirically confirmed the supposition. In our case, only 30% of our 699 instances have one of cue phrase markers shown in Table 4, though this value can be dependent of the data.

This result suggests that in order to develop knowledge acquisition methods for causal relations with high coverage, we must deal with linguistic expressions with no explicit cue phrase markers as well as those with cue phrase markers.

7.2 The parts-of-speech of head elements

Next, we classified the events included in the 699 instances into two syntactic categories: the verb phrase (VP) and the noun phrase (NP). To do this, we used morphological information of their head elements. If the part-of-speech of a head is verb or adjective, the event is classified as a verb phrase. If

---

[3]The cue phrase markers whose frequencies are less than three are not listed due to space limitation in Table 4.

Table 5: The syntactic types

		e_1	e_2
VP	(verb) (adjective)	365	412
NP	(verbal noun) (general noun)	322	269
others		12	18

the part-of-speech of a head is noun (including general noun and verbal noun), the event is classified as a noun phrase. We used *ChaSen* [4] to get part-of-speech information.

The result is shown in Table 5. More than half events are classified as the VP. This matches our intuition. However, the number of events classified as the NP is comparable to the number of events classified as the VP; 322 events of e_1 are represented as noun phrases, and 269 events of e_2 are also represented as noun phrases.

This result is quite suggestive. To promote the current methods for knowledge acquisition to further stage, we should develop a knowledge acquisition framework applicable both to the verb phrases and to the noun phrases.

7.3 The positions of head elements

For each e_1 and e_2 included in the 699 instances, we examined the positions of their head elements in the sentences.

We consider dependency structures between *bunsetsu*-phrases in the original sentences from which causal relation instances are extracted. The dependency structures form tree structures. The *bunsetsu*-phrase located in the end of the sentence is the root node of the tree. We focus on the depth of the head element from the root node. We used *CaboCha*[5] to get dependency structure information between *bunsetsu*-phrases.

The results are shown in Figure 3 and Figure 4. Figure 3 is the result for the head elements of e_1, and Figure 4 is the result for the head elements of e_2. The letter "f" in Figure 3 and Figure 4 indicates frequency at each position. Similarly, the letter "c"
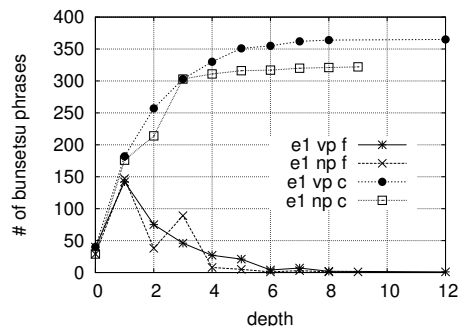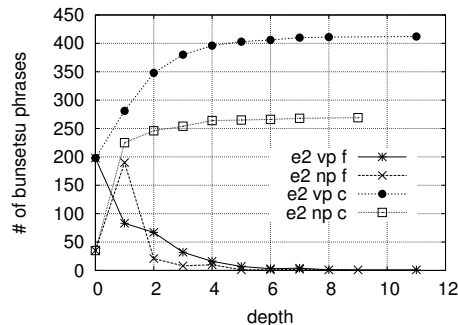
Figure 3: The positions of head elements (e_1)

Figure 4: The positions of head elements (e_2)

indicates cumulative frequency.

In Figure 4, the 198 head elements of the events represented as a verb phrase are located in the end of the sentences, namely depth = 0. The 190 of the 269 events represented as a noun phrase are located in depth = 1. For events represented as either a verb phrase or a noun phrase, over 80% of head elements of the events are located within depth < 3. In Figure 3, similarly, over 80% of head elements of the events are located within depth < 4.

These findings suggest that the most of the events are able to be found simply by searching the *bunsetsu*-phrases located in the shallow position at the phase of causal knowledge acquisition.

7.4 Relative positions of two head elements

Finally, we examined relative positions between head elements of e_1 and e_2 where these two events are held in a causal relation. In Section 7.3, we discussed each absolute position for e_1 and e_2 by means of the notion of depth in sentences. Here, we focus on the difference (D) of the depth values between e_1 and e_2.

The result is shown in Table 6. The symbol "$e_1 \Rightarrow e_2$" in Table 6 indicates the case where the head element of e_1 is located nearer to the beginning of the

Table 6: Relative positions of two head elements

		$e_1 \Rightarrow e_2$	$e_2 \Rightarrow e_1$
intra-sentential	$D = 1$	259	15
	$= 2$	152	23
	> 2	33	4
	no dep	72	
inter-sentential		141	

sentence than that of e_2. The "$e_2 \Rightarrow e_1$" indicates the opposite case. The symbol "no dep" indicates the case where neither the condition *a* nor *b* is satisfied:

 a. the head element of e_2 is an ancestor of the head element of e_1.

 b. the head element of e_2 is a descendant of the head element of e_1.

The symbol "inter-sentential" indicates the case where two head elements appear in different sentences.

 The most instances（259 instances）are categorized into $D = 1$ on $e_1 \Rightarrow e_2$, that is, the head element of e_1 directly depends on the head element of e_2. This result matches our intuition. However, there are several other cases. For example, 152 instances are categorized into $D = 2$ on $e_1 \Rightarrow e_2$, 72 instances are categorized into "no dep". Most of the instances extracted from sentences including any parallel relations are categorized into "no dep". In this study, we consider causal relation instances as binary relation. To deal with instances categorized into "no dep" adequately, we should extend our framework to the more complex structure.

8 Conclusion

We reported our causal relation tags and the annotation workflow. Using the annotated corpus, we examined the causal relation instances in Japanese text. From our investigation, it became clear that what amount of causal relation instances are present, where these relation instances are present, and which types of linguistic expressions are used for expressing these relation instances in text.

Acknowledgement

References

B. Altenberg. 1984. Causal linking in spoken and written English. *Studia Linguistica*, 38:1.

R. Girju. 2003. Automatic detection of causal relations for question answering. In *Proc. of the 41st ACL, Workshop on Multilingual Summarization and Question Answering*.

T. Inui. 2004. *Acquiring causal knowledge from text using connective markers*. Ph.D. thesis, Graduate School of Information Science, Nara Institute of Science and Technology.

Y. Liu. 2004. *Semantics and usages of connectives for causal relations in modern Japanese - cases of 'dakara', 'sitagatte', 'soreyue(ni)', 'sonokekka', 'sonotame(ni)' -*. Ph.D. thesis, The Graduate School of Languages and Cultures, Nagoya University.

Mainichi. 1995. Mainichi Shimbun CD-ROM version.

D. Marcu. 1997. *The rhetorical parsing, summarization, and generation of natural language texts*. Ph.D. thesis, Department of Computer Science, University of Toronto.

E. Miltsakaki, R. Prasad, A. Joshi, and B. Webber. 2004. Annotating discourse connectives and their arguments. In *Proc. of the HLT/NAACL Workshop on Frontiers in Corpus Annotation*.

S. Nishizawa and Y. Nakagawa. 1997. A method of discourse structure understanding in Japanese task-free conversation for causal conjuction. *Natural Language Processing*, 4(4):61–72.（in Japanese）.

M. Palmer, D. Gildea, and P. Kingsbury. 2005. The proposition bank: A corpus annotated with semantic roles. *Computational Linguistics Journal*, 31(1).

J. Pustejovsky, J. M. Castaño, R. Ingria, R. Sauri, R. J. Gaizauskas, A.Setzer, G. Katz, and D. R. Radev. 2003. TimeML: Robust specification of event and temporal expressions in text. In *New Directions in Question Answering*, pages 28–34.

H. Sato, K. Kasahara, and K. Matsuzawa. 1999. Rertrieval [sic] of simplified causal knowledge in text and its application. In *Technical report of IEICE, Thought and Language*.（in Japanese）.