

Toward a Discourse Theory for Annotating Causal Relations in Japanese

Kimi Kaneko
Ochanomizu University
kaneko.kimi@is.ocha.ac.jp

Daisuke Bekki
Ochanomizu University
National Institute of Informatics
CREST, JST
bekki@is.ocha.ac.jp

Abstract

We present a revised discourse theory based on segmented discourse representation theory and provide a method for building a Japanese corpus suitable for causal relation extraction. This extends and refines the framework proposed in Kaneko and Bekki (2014), and we evaluate our corpus and compare it with that work.

1 Introduction

In recent years, considerable attention has been paid to deep semantic processing. Many studies, including Bethard et al. (2008), Inui et al. (2007), Inui et al. (2003), and Riaz and Girju (2013), have recently been conducted on deep semantic processing, and causal relation extraction (hereinafter, CRE) is one of the specific tasks of deep semantic processing. Research on CRE is still progressing, and there are many obstacles that must be overcome.

In Inui et al. (2003), cause and effect pairs were acquired from Japanese texts by using keywords such as “node” and “kara”. In (1), for example, the antecedent *ame-ga hut-ta* (“it rained”) denotes an event taken as a cause, and the consequent *mizutamari-ga dekita* (“puddles emerged”) denotes an event taken as an effect.

- (1) Ame-ga hut-ta-node
rain-NOM fall-past-because
mizutamari-ga deki-ta.
puddles-NOM emerge-past
‘Because it rained, puddles emerged.’

However, antecedents do not always denote causes or reasons for consequents, as illustrated by the following example.

- (2) Kesa kubi-ga
this.morning neck-NOM
itakat-ta-node
have.a.pain-past-because
netigae-ta-no daroo.
strain.my.neck-past-attr may.
‘Because I had a pain in my neck this morning, I might have strained my neck while sleeping.’

In example (2), the antecedent *kesa kubi-ga itakat-ta* (“I had a pain in my neck this morning”) is not taken as the cause of the consequent *netigae-ta* (“I strained my neck while sleeping”) but as the basis for the judgment expressed by the consequent. In this example, the consequent denotes the cause, and the antecedent denotes its effect. For a computer to automatically recognize causal relations in text, it is important to distinguish cases like (2) from cases like (1). However, existing studies have not dealt with these kinds of problems.

To solve such problems, Kaneko and Bekki (2014) (henceforth K&B) analyzed the information necessary for acquiring more accurate cause–effect knowledge and proposed a method for creating a Japanese corpus suitable for CRE. However, as is explained below, some problems remain: first, the coverage of discourse relations is not sufficient; second, annotating at two levels (i.e. fact and epistemic levels) is sometimes redundant.

We try to solve these remaining problems in K&B; toward that end, we propose a new method for building a Japanese corpus for causal relation extraction. In addition, we evaluate the validity of our method in terms of agreement and frequency, and analyze the results.

2 Previous Studies

In this section, we introduce some of the previous studies on annotation of temporal, causal, and other relations, as well as some linguistic analyses of temporal, causal and discourse relations.

Bethard et al. (2008) generated English data sets annotated with temporal and causal relations and analyzed interactions between the relations. In addition, these specialized data sets were evaluated in terms of inter-annotators agreement and accuracy. Relations were classified into two causal categories (CAUSAL, NO-REL) and three temporal categories (BEFORE, AFTER, NO-REL). With regard to the evaluation, they pointed out that the classification was coarse and that reanalysis with finer relations would be necessary. Moreover, they reported that some event pairs have ambiguous temporal relations. For (3), for example, it was difficult for most annotators to judge which event of “was ahead from the start” and “don’t need to invite in competitive allies” precedes the other and how much the events temporally overlap.

- (3) IBM established its standard to try to stop falling behind upstart Apple Computer, but NEC [*EVENT was*] ahead from the start and didn’t [*EVENT need*] to invite in competitive allies.¹

Inui et al. (2005) characterized causal expressions in Japanese text and built a Japanese corpus with tagged causal relations. However, usages such as that illustrated in (2) and interactions between temporal relations and causal relations were not analyzed.

Asher and Lascaridas (2003)’s segmented discourse representation theory (SDRT) is a formal discourse theory that accounts for cases in

¹This sentence was extracted from Bethard et al. (2008).

which discourse relations (rhetorical relations) interact with the truth-conditional meanings of sentences. Some of the discourse relations in SDRT have constraints on temporal and causal relations, so that we can calculate semantic contents that interact with them by means of sequences of logical reasoning. Consequently, we can build a corpus for CRE in which we have considered the influences of discourse and temporal relations by annotating not only causal relations but also discourse relations in SDRT into text. As examples of theories of discourse relations, we mention especially rhetorical structure theory (RST) (Mann and Thompson, 1987) and cross-document structure theory (CST) (Radev, 2000). One of the problems that they equally share is the inability to exhibit sequences of reasoning based on nonverbal information for specifying discourse relations. To solve this problem, exhibiting a process of sequences of reasoning should be possible.

K&B reframed SDRT by distinguishing between discourse relations, temporal relations, and causal relations, and annotated Japanese texts with these three types of relations. These relations are assigned at two levels: the fact level, which describes the fact that is actually occurring in the real world, and the epistemic level, which describes what the speaker recognizes as the fact. Example (1) is annotated as in (4).

- (4) Fact-level: [**Precedence**(π_1, π_3),
Explanation(π_1, π_3),
CAUSE(π_1, π_3)],
Epistemic-level: [**Precedence**(π_2, π_4),
Explanation(π_2, π_4),
CAUSE(π_2, π_4)],
 $\pi_2 \pi_1$ Ame-ga hut-ta-node,
 $\pi_4 \pi_3$ mizutamari-ga dekita.

According to K&B, discourse relations and causal relations impose some restrictions on the interpretation of temporal relations. For example, the relation CAUSE(A,B) imposes the temporal relation Precedence(A,B). In (4), CAUSE(π_1, π_3) imposes the temporal constraint Precedence(π_1, π_3) at the fact level; at the same time, CAUSE(π_2, π_4) imposes

Precedence(π_2, π_4) at the epistemic level. In the following example, by contrast, the causal and temporal relations at the fact level that hold between the main clause and the subordinate clause are reversed at the epistemic level.

- (5) Fact-level: [**Precedence**(π_3, π_1),
Explanation(π_1, π_3),
CAUSE(π_3, π_1)],
 Epistemic-level: [**Precedence**(π_2, π_4),
Explanation(π_2, π_4),
CAUSE(π_2, π_4)],
 $\pi_2 \pi_1$ Kesa kubi-ga itakat-ta -*node*,
 $\pi_4 \pi_3$ netigae-ta-no-daroo.

Note that by distinguishing the two levels, the temporal constraints that discourse and causal relations impose are kept consistent. In (5), the temporal constraints Precedence(π_3, π_1) and Precedence(π_2, π_4) hold at different levels, so no contradiction arises here. In this way, K&B can handle examples, like (5), that involve an apparent mismatch between causal and temporal relations.

However, it is not clear whether this approach would also be effective for a large-scale corpus because the data sets built by K&B are relatively small. In this study, we follow the approach of K&B and attempt to build a Japanese corpus tagged with discourse relations for CRE. In the course of doing so, we have discovered that the theory of K&B has the following problems.

- The coverage of discourse relations is sometimes insufficient. The balanced corpus of contemporary written Japanese (BCCWJ) (Maekawa, 2008) is designed as a corpus that contains texts of various styles, and further annotation has revealed that the set of discourse relations in K&B covers only some parts of the possible relations.
- Annotating both fact- and epistemic-level information for every pair of segments is redundant. As mentioned above, the distinction between fact- and epistemic-level information plays an important role in K&B, but in most cases the information will coincide.

- Judging the temporal relation between the events is not an easy task; however, the “Narration” family can only be further classified by means of the temporal relations. We can highlight the difficulty by the following simple example (6).

- (6) a. Nippon-no natu-wa atui.
 Japan-GEN summer-TOP be.hot
 Ippou-de, Nippon-no
 on.the.other.hand Japan-GEN
 huyu-wa samui.
 winter-TOP be.cold
 ‘The summer is hot in Japan. On the other hand, the winter is cold in Japan.’

- b. Fact-level: [**Narration**(π_1, π_3)?,
Overlap(π_1, π_3)?]²,
 Epistemic-level: [**Narration**(π_2, π_4)?,
Overlap(π_2, π_4)?],
 $\pi_2 \pi_1$ Nippon-no natu-wa atui.
 π_4 Ippoo-de, π_3 Nippon-no huyu-wa samui.

One may tag this sentence pair with the “Narration” label, which actually includes different kinds of narrative relations, such as “Background” and “Parallel” relations, depending on the temporal relation between them. However, both sentences of example (6) are generic, which means there is no temporal order between the things described. In other words, we can decide neither which fact occurred earlier nor which fact was recognized earlier.

One may additionally tag this sentence with the “Overlap” label described in K&B, but it is apparent that the times spanned by “summer” never overlap those spanned by “winter,” contradictory to what the “Overlap” label means. Because of this, the descriptive power of the label set described in K&B is insufficient, and so we have to reconsider the label system.

This study aims to rearrange K&B’s theory on the basis of further reflections on SDRT as a means to rebuild an exhaustive theory of discourse relations for CRE. First, we propose

²In K&B, pairs of sentences are not tagged with causal relations when there is no causal relation.

a new annotation scheme to solve the above-mentioned problems. Second, we focus on the first problem and annotate sentences with this new setting. Finally, we evaluate and analyze our annotation scheme and the data set.

3 Method

We extended and refined K&B and developed a new method for annotating the relation between the following segment pairs in discourse.

1. A discourse and a subsequent sentence
2. A main clause and its subordinate clause
 - (a) when the predicate of subordinate clause is in continuation form or “te”-form)
 - (b) when two clauses are connected by causal suffixes (e.g. “node”, “kara”)

3.1 Causal Relation

We distinguish two different kinds of ‘causal’ relations and annotate them separately. **Explanation** is a discourse relation, which is a relation between two linguistic expressions, and **Cause** is a causal relation between two propositions.³

As a discourse relation, **Explanation** is a relation between two adjacent segments: in other words, it is a grammatical relation between two constituents. In contrast, **Cause** is a relation between facts and not restricted to two adjacent segments. **Cause**(A,B) is the only causal relation that we adopt, as shown in Table 1. We use the tag only when there exists a causal relation between a pair of propositions A and B; in other cases, no annotation is used.

The distinction between these two relations is essential because, on one hand, the causality may not be expressed linguistically and, on the other hand, a linguistically claimed causal relation does not ensure actual causality.

As an example of the former case, consider (7), where John’s putting the banana peel is in reality a possible cause of Bill’s tumbling even

³A “proposition” in this paper is a tensed predicate (e.g. “fall,” “have a pain,” and “strain” in (1)(2)) whose eventuality is either an event or a state, along with its arguments and modifiers.

though it is not linguistically marked. The discourse relation between the two sentences here is **Narration**, which specifies only a temporal relation between them (as consecutive events).

- (7) John-ga banananokawa-wo
John-NOM banana.peel-ACC
yuka-ni oi-ta. Bill-ga
the.floor-LOC put-past Bill-NOM
koron-da.
tumble-past
‘John put a banana peel on the floor.
Then, Bill tumbled.’

The latter case is exemplified by (8), where the speaker claims that John’s rain-making ritual caused the rain, by using a causal discourse relation, although nobody can tell whether it is in fact so.

- (8) John-ga amagoi-wo
John-NOM rain.making.ritual-ACC
si-ta-node ame-ga fut-ta
do-past-cause rain-NOM fall-past
noda.
epistemic.modal-pres
‘Because John performed a rain-making
ritual, it rained.’

3.2 Without a Fact/Epistemic Level Distinction

Unlike K&B, we abolish the distinction between the factual and epistemic levels. We distinguish only whether a given segment is propositional or modal (including the segment with the causal suffix “node” and the epistemic modal suffix “noda”: the former roughly corresponds to the fact level, and the latter to the epistemic level).

This decision substantially simplifies and reduces the work of annotators. However, the distinction between fact and epistemic levels is one of the core ideas in K&B used to avoid temporal contradiction, which we described in Section 1. Therefore, we have to show how our simplified setting is still free from that contradiction.

As examples, the result of annotating (1) and (2) are shown in (9) and (10), respectively.

Level	Description
Cause(A,B)	The proposition A is a cause of the proposition B.

Table 1: Causal relation

- (9) [**Explanation**(π_2, π_3), **Cause**(π_1, π_3)]
 $\pi_2 \pi_1$ Ame-ga hut-ta-node,
 π_3 mizutamari-ga dekita.
 a”. Temporal relation:
 Precedence(π_1, π_3), Precedence(π_2, π_3)

- (10) [**Explanation**(π_2, π_4), **Cause**(π_3, π_1)]
 $\pi_2 \pi_1$ Kesa kubi-ga itakat-ta -node,
 $\pi_4 \pi_3$ netigae-ta-no-daroo.
 a”. Temporal relation:
 Precedence(π_3, π_1), Precedence(π_2, π_4)

Both **Cause(A,B)** and **Explanation(A,B)** require that a temporal relation Precedence(A,B) holds⁴ since a cause must precede its effect (otherwise, it is not a cause-effect relation). The issue is determining whether these two relations impose contradictory temporal relations.

In (10), the antecedent part π_2 of the conditional has the causal suffix “node”, which embeds a propositional part π_1 , and the consequent part π_4 has the modal suffix *daroo*, which also embeds a propositional part π_3 .

Because **Explanation** is a discourse relation, it is a relation between a pair of adjacent segments π_2 and π_4 . As a result, it is a relation between two modal expressions, stating that “Realizing that I had a pain in my neck caused me to infer that I strained my neck,” which is as expected. The temporal requirement Precedence(π_2, π_4) is that realization of a pain precedes the inference of the strain, which is also as expected.

In contrast, **Cause**(π_3, π_1) in (10) is a causal relation between the propositions π_3 and π_1 , namely, straining the neck is a cause of the pain. Here, the temporal requirement is Precedence(π_3, π_1), which states that the strain must precede the pain.

⁴For the sake of brevity, we do not discuss the details of temporal relations in this paper.

In this way, two different precedence relations, one at the fact level and the other at the epistemic level, can be properly treated in this setting, without introducing fact and epistemic levels to every segment.

3.3 Discourse Relations

In addition to **Explanation**, we have the set of discourse relations, based on SDRT and K&B, shown in Table 2. It is also shown in Table 3 how discourse relations in our method correspond to those in K&B and those in SDRT. As Table 3 displays, discourse relations in our study integrate the temporal relations and discourse relations of K&B.

Moreover, a procedure to identify discourse relations in our method is shown below.

Procedure:

1. First, judge the logical relation between the pair A and B to determine whether it is conjunctive, disjunctive, or conditional (by the standard truth-conditional tests):
 - (a) If it is disjunctive, tag it with the **Alternation** label.
 - (b) If it is conditional, tag it with the **Consequence** label.
 - (c) If it is conjunctive, proceed to 2.
2. Judge whether the relation is adversative or contrastive:
 - (a) If it is adversative, proceed to 3.
 - (b) If it is contrastive, especially when expressions such as “*sikashi*”, “*tokoroga*” appear, tag it with the **Contrast** label.
3. (a) If B describes an event that is a part of the whole event described by A, then tag the relation with the **Elaboration** label.

⁵Temp-rel(A,B) \equiv Precedence(A,B) \vee Overlap(A,B) \vee Subsumption(A,B)

Label	Description
Alternation (A,B)	“A or B”: logical disjunction ($A \vee B$).
Consequence (A,B)	“If A then B”: logical implication ($A \rightarrow B$).
Contrast (A,B)	“A but B”: B contrasts with A.
Elaboration (A,B)	B describes a part of A in detail.
Explanation (A,B)	A is a cause and B is its effect.
Commentary (A,B)	The content of A is summarized or complemented by B.
Instance (A,B)	“A, for example, B’,” where B describes an instance of A.
Addition (A,B)	The description of the state B is added to the description of the state A.
Parallel (A,B)	The two events A and B overlap.
Narration (A,B)	The occurrence of the event B is subsequent to that of A.
Introduction (A,B)	B introduces a new reference point that is independent from that of A.
Background (A,B)	B describes the background situation of the event A.

Table 2: Discourse relation list

Ours	SDRT	K&B
Cause (A,B)	Explanation(A,B)	CAUSE(A,B)
Alternation (A,B)	Alternation(A,B)	Alternation(A,B)
Consequence (A,B)	Consequence(A,B)	Consequence(A,B)
Contrast (A,B)	Contrast(A,B)	Contrast(A,B)
Elaboration (A,B)	Elaboration(A,B)	Elaboration(A,B)
Explanation (A,B)	Result(A,B)	Explanation(A,B)
Commentary (A,B)	Commentary(A,B)	Commentary(A,B)
Instance (A,B)	–	–
Addition (A,B)	Parallel(A,B)	Narration(A,B) \wedge Overlap (A,B)
Parallel (A,B)		Parallel(A,B)
Narration (A,B)	Narration(A,B)	Narration(A,B) \wedge Precedence (A,B)
Introduction (A,B)	Narration(A,B)	Narration(A,B) \wedge Temp_rel(A,B) ⁵
Background (A,B)	Background(A,B)	Narration(A,B) \wedge Subsumption (A,B)

Table 3: Correspondence among K&B, SDRT, and our method

- (b) If A describes the basis of the judgment B, particularly indicated when an expression such as “*dakara*”, “*sitagatte*”, or “*yueni*” appears, tag the relation with the **Explanation** label.
 - (c) If B is a summary, a restatement, or a complementary remark of A, especially when expressions such as “*tumari*” and “*yoosuruni*” appear, then tag the relation with the **Commentary** label.
 - (d) If A is a universal or generic sentence and B is an instance of A, then tag the relation with the **Instance** label.
 - (e) Otherwise, proceed to 4.
4. Judge whether the eventualities of A and B are events or states, and the reference points (in the sense of tense) of A and B:
- (a) If both A and B are states, then tag the relation with the **Addition** label.
 - (b) If both A and B are events and they take place in the same time span (and overlap), then tag the relation with the **Parallel** label.
 - (c) If B is an event and B happens successively to A, then tag the relation with the **Narration** label. Specifically, do so when B’s reference point is just after A’s reference point.
 - (d) If A is a state, B is an event, and B introduces a reference point that is independent of A’s reference point, then tag the relation with the **Introduction** label.
 - (e) If A is an event, B is a state, and B’s reference point is the same as A’s, then tag the relation with **Background** label.

The whole decision process is depicted by Figure 1.

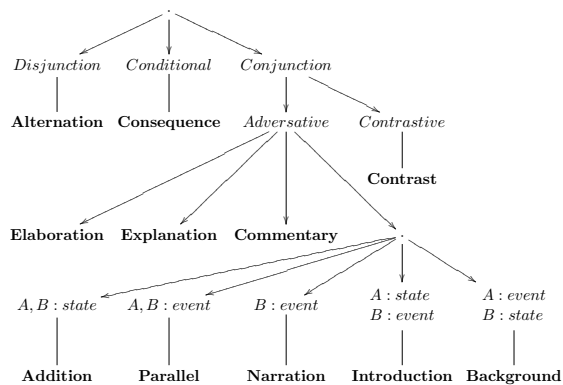


Figure 1: Decision tree for discourse relations

3.4 Comparison with Kaneko and Bekki (2014)

We compare our approach to that of K&B.

First, we refined the set of discourse relations in K&B by adding new discourse relations, as necessitated by the cases that K&B’s set of discourse relation does not cover.

Second, we proposed a decision procedure in Figure 1 for classifying a discourse relation. We consider this as a substantial advance since the previous criteria for identifying discourse relations in K&B are vague and we believe that we could make them clearer.

Third, we abolished the distinction between the fact and epistemic levels, which makes our annotation far simpler than that of K&B, while still enabling us to deal with the cases, such as (2), in which the temporal precedence relations at the fact and epistemic levels seem to contradict, as discussed in Section 3.

4 Results

We applied our method to 128 sentences from BCCWJ (Maekawa, 2008). The labels were assigned to the sentences by two annotators. During labeling, we used the labels presented in Section 3. Our method was developed on the basis of 73 sentences, and by using the 73 sentences and the other 55 sentences, we evaluated

Label	K&B			Ours (sentences)
	Total	Fact	Epistemic	Total
Precedence	25	14	11	–
Overlap	7	4	3	–
Subsumption	61	29	32	–
Total	94	47	47	–
CAUSE	14	8	6	6
Total	14	8	6	6
Alternation	–	–	–	–
Consequence	6	3	3	–
Explanation	14	7	7	9
Contrast	2	1	1	6
Commentary	–	–	–	6
Narration	66	33	33	52
Background	–	–	–	1
Addition	–	–	–	17
Parallel	–	–	–	0
Introduction	–	–	–	8
Elaboration	4	2	2	23
Instance	–	–	–	6
Total	94	47	47	128

Table 4: Distribution of labels to segments in our study for the BCCWJ (italicized labels are newly added).

the inter-annotators agreement and kappa coefficient as well as the number of annotations and compared the results with those of K&B. The agreement for 128 sentences was 0.67 and was computed as follows (the kappa coefficient was 0.57).

$$Agreement = \text{Identical labels} / \text{Total labels}$$

K&B reported an agreement rate of 0.68, although they computed the agreement by using annotated segment data, which means the results are not directly comparable to ours. Nevertheless, the close values suggest that our method is comparable to that in K&B’s study in terms of agreement.

Analyzing more segments in actual text and improving our method could lead to further improvement in terms of agreement.

An average of 20 and 11 sentences were tagged per hour in our study and K&B’s study, respectively. This indicates that the complexity of our method is not much different from that in K&B.

Table 4 shows the distribution of labels into segments in K&B and into sentences in our study. **Background**, **Addition**, **Parallel**, **Introduction** and **Instance** were newly added in our study. “Narration” in K&B covers **Back-**

ground, **Addition**, and **Parallel**. “Elaboration” in K&B and SDRT includes **Instance**. While **Narration** was the most frequently used label, and so it was biased greatly in K&B, the frequency of each relation in our study is more balanced than that in K&B. Thus, the classification in our study is more appropriate for performing machine learning. However, whether our method is truly more appropriate than K&B should be judged by annotating segments with our relations and comparing those results with K&B.

We can see from Table 4 that **Narration** was still the most frequently used label, and some labels, such as **Alternation**, never appeared. As a result, we can assume that frequent relations will be distinct from non-frequent relations. So far, all relations are either frequent or non-frequent, although a larger data set should be analyzed to confirm this.

5 Discussion

We analyzed errors in this annotation exercise. First, under the current version of our annotation guideline, some judgments inevitably remain ambiguous. We explain when and why this happens, in a mini-discourse example shown in Fig.2 (p.10). The annotators do not agree with the results of annotations for $\pi 4$ in the sentence (14): their results range over **Addition**($\pi 1, \pi 4$) (or **Addition**($\pi 2, \pi 4$)), **Commentary**($\pi 1, \pi 4$), and **Narration**($\pi 3, \pi 4$). The problem is that this case may be actually ambiguous among these three cases, and none of the choices alone adequately explains the discourse relation.

The **Addition** label here breaks the continuous structure from $\pi 1$ to $\pi 4$ by skipping $\pi 3$ and directly connects to $\pi 1$ or $\pi 2$, which is due to the restriction that its first argument must be a state but $\pi 3$ is an event. However, this choice does not correctly capture the structure of the discourse, in which the sequence of sentences $\pi 1$ to $\pi 4$ seem to incrementally add information to the discourse.

The **Commentary** label currently covers several heterogeneous cases: (1) the case that

the second argument is a summary of the first argument, (2) the case that the second argument is a restatement of the first argument, and (3) the case that the second argument adds some supplementary comments (such as footnotes). Now, the third case applies to $\pi 4$; however, this is not distinguishable from the **Addition** label. It is necessary to separate the different uses of this label.

The **Narration** label, which has a restriction that the second argument must be a state, is reasonable if we assume that the verb with the aspect “teiru” (a perfect suffix) in $\pi 4$ denotes an event. However, the reference time of $\pi 3$ is in the year 2004 while that of $\pi 4$ is a speech time, which is a bit too long of a time span to consider $\pi 3$ and $\pi 4$ to be sequential.

Second, there are problems in annotating non-assertive sentences, such as interjections, exclamatory sentences, and rhetorical questions. They appear not only in dialogue but also in monologue, which causes difficulty in making a judgment about their discourse relation to previous sentences.

At present, we treat interjections such as “Ooops!” as if they are ellipses: for example, we may regard its full form as “I cried out ooops!” and judge their relations accordingly.

The cases of rhetorical questions such as “how do I know it?”, as in 15, can be treated in the same way; for example, here we regard it as an elliptical form of the full form “I wonder how I know it.”

- (15) Dare-demo sorekurai-wa
 everyone-NOM to.such.extent-ACC
 taiken-siteiru daroo to
 experience-perf may that
 souzou-siteiru noda-ga,
 imagine-prog I.know-but
 doudaroo.
 how.do.I.know

‘I imagine that everyone may have experienced such things, but how do I know?’

However, it should be further investigated whether this method can be applied to all cases in a uniform and principled manner.

6 Conclusions

We proposed a method for discourse annotation based on a discourse theory that revises and extends that of K&B as a means of building a more precise Japanese corpus for CRE. We have annotated 128 sentences in BCCWJ with discourse relations and causal relations, and compared the annotations of 128 of these sentences with the annotations in K&B in terms of agreement, kappa coefficient, frequencies, and time needed for decomposition. We reported and analyzed the results and discussed some problems of our method. For future work, we intend to address the problems we described in Sections 4 and 5 by the further refinement of our discourse theory.

Acknowledgments

Our sincere thanks to Koji Mineshima and Maika Utsugi for many helpful discussions, and to the anonymous reviewers of PACLIC28 for insightful comments. This research is partially supported by JST, CREST and Data Centric Science Research Commons.

References

- Asher N. and Lascaridas A. 2003. *Logics of Conversation: Studies in Natural Language Processing*. Cambridge University Press, Cambridge, UK.
- Bethard S., Corvey W. and Kilingerstein S. 2008. *Building a Corpus of Temporal Causal Structure*. LREC 2008, Marrakech, Morocco.
- Inui T., Inui K. and Matsumoto Y. 2005. *Acquiring Causal Knowledge from Text Using the Connective Marker Tame*. ACM Transactions on Asian Language Information Processing (ACM-TALIP), Vol.4, Issue 4, Special Issue on Recent Advances in Information Processing and Access for Japanese, 435–474.
- Inui T., Inui K. and Matsumoto Y. 2003. *What Kinds and Amounts of Causal Knowledge Can Be Acquired from Text by Using Connective Markers as Clues*. The 6th International Conference on Discovery Science (DS-2003), 180–193.
- Inui T., Takamura H. and Okumura M. 2007. *Latent Variable Models for Causal Knowledge Acquisition*. Alexander Gelbukh(Ed.), *Computational Linguistics and Intelligent Text Processing, Lecture Notes in Computer Science*, 4393:85–96.
- Kaneko K. and Bekki D. 2014. *Building a Japanese Corpus of Temporal-Causal-Discourse Structures Based on SDRT for Extracting Causal Relations*. EACL-2014 Workshop on Computational Approaches to Causality in Language, 33–39.
- Maekawa K. 2008. *Balanced Corpus of Contemporary Written Japanese*. In Proceedings of the 6th Workshop on Asian Language Resources (ALR), 101–102.
- Mann W. C. and Thompson S. 1987. *Rhetorical Structure Theory: A Theory of Text Organization*. ISI Reprint Series, ISI/RS-87-190, 1–82.
- Radev D.R. 2000. *A common theory of information fusion from multiple text sources step one: cross-document structure*. In SIGDIAL '00 Proceedings of the 1st SIGdial workshop on Discourse and dialogue, Volume 10, 74–83.
- Riaz M. and Girju R. 2013. *Toward a Better Understanding of Causality between Verbal Events: Extraction and Analysis of the Causal Power of Verb-Verb Associations*. In Proceedings of the SIGDIAL 2013 Conference, Metz, France 21–30.

- (11) π 1:Bunkatyoo-bunkakooryuusi-zigyow-wa, nipponbunka-ni
 Agency.for.Cultural.Affairs's.cultural.ambassador.project-TOP Japanese.culture-DAT
 tazusawaru hitobito-ni "Bunkakooryuusi"-tosite nipponbunka-wo
 be.concerned.with-attr people-DAT as."cultural.ambassadors" Japanese.culture-ACC
 hiromete-moraukoto-wo mokuteki-tosite 2003-nendo-kara hazimeta zigyow-desu.
 to.make.promote-ACC aim.for-cont from.year.2003 start-attr be.project-past
 'The cultural ambassador project of the Agency for Cultural Affairs is a project that started
 in the year 2003 with the aim of appointing people who are concerned with Japanese culture
 to promote Japanese culture, as "cultural ambassadors."'
- (12) π 2:"Bunkakooryuusi"-no-katudoo-niwa, (i) "kaigaihaken-gata",
 "cultural.ambassador"'s.activity-TOP "overseas.dispatching-type"
 (ii) "gentitaizaisya-gata", (iii) "rainitigeizyutuka-gata"-no 3tu-no-ruikei-ga aru.
 "immigrant-type" "visiting.artist-type" 3.types-NOM exist-pres
 'There are three types of "cultural ambassador" activities: (i) "the overseas dispatching
 type," (ii) "the immigrant type," and (iii) "the visiting artist type".'
- (13) π 3:2004-nendo-wa, "kaigaihaken-gata"-bunkakooryuusi-tosite 11-me,
 In.year.2004-TOP as."overseas.dispatching-type"-cultural.ambassador 11-people
 "gentitaizaisya-gata"-bunkakooryuusi-tosite 4-me,
 as."immigrant-type"-cultural.ambassador 4-people
 "rainitigeizyutuka-gata"-tosite 4-kumi-no sime-wo
 as."visiting.artist-type"-cultural.ambassador" 4-teams-ACC appoint-ACC
 okonai-masi-ta.
 execute-polite-past
 'In the year 2004, the Agency for Cultural Affairs appointed 11 people as "overseas dis-
 patching type" cultural ambassadors, 4 people as "immigrant type" cultural ambassadors,
 and 4 teams as "the type of artist who visits Japan" cultural ambassadors.'
- (14) π 4:Nipponbunka-ni nazimino-usukat-ta kuni ya tiiki-deno nipponbunnka-no
 Japanese.culture-DAT unfamiliar.with-past country and area-LOC Japanese.culture-ACC
 syoukai-wo okonatte-i-masu.
 introduce-ACC execute-prog-polite
 'Cultural ambassadors are introducing Japanese culture in countries and areas that were
 unfamiliar with Japanese culture.'

Figure 2: Example sentences and annotations