Recognition of the Coherence Relation between Te-linked Clauses

Akira Oishi
School of Information Science
JAIST
1-1 Asahidai, Tatsunokuchi,
Ishikawa 923-1292, Japan
oishi@jaist.ac.jp

Yuji Matsumoto
Graduate School of Information Science
NAIST
8916-5 Takayama, Ikoma,
Nara 630-0101, Japan
matsu@is.aist-nara.ac.jp

Abstract
This paper describes a method for recognizing coherence relations between clauses which are linked by te in Japanese — a translational equivalent of English and. We consider that the coherence relations are categories each of which has a prototype structure as well as the relationships among them. By utilizing this organization of the relations, we can infer an appropriate relation from the semantic structures of the clauses between which that relation holds. We carried out an experiment and obtained the correct recognition ratio of 82% for the 280 sentences.

1 Introduction
One of the basic requirements for understanding discourse is recognizing how each clause coheres with its predecessor. Our linguistic and pragmatic competence enables us to read in conceivable relations even when two clauses are copresent without any overt cues, i.e., in parataxis.

There has been a variety of definitions for coherence relations (see (Hovy and Maier, 1993) for a survey). However, the definitions are rather vague and they are often recognized to be underspecified (Moore and Pollack, 1992; Fukumoto and Tsujii, 1994). This paper attempts to explicate how such coherence relations arise between segments of discourse. We focus on te-linkage in Japanese — a translational equivalent of English and-linkage, since mere parataxis ranges over too widely to capture the underlying principles on the coherence relations.

We consider that coherence relations are categories each of which has its prototypical instances and marginal ones. As with all instances of categorizations, the prototypical cases of each relation are clearly distinguishable from one another. In some cases, however, it is often hard to make clear argument for a relation being one rather than another. In addition, these relations themselves are hierarchically organized according to their specificity. By considering the prototype of each relation, we can infer an appropriate relation from the semantic structures of the segments between which that relation holds.

2 Categorization of Te-linkage
Traditionally, te-constructions have been divided into three categories according to the function of te: (i) as a non-productive derivational suffix; (ii) as a linker joining a main verb with a so-called auxiliary to form a complex predicate; and (iii) as a linker connecting two phrases or clauses. Since the derivatives and the auxiliaries are relatively fixed compared with the third category, we concentrate on the third category in this paper.

Japanese te, like English and, is used to express a diverse range of coherence relations as shown below.

(1) Circumstance
itami-wo korafe hasiri-tuzuketa.
pain-ACC endure-te run-continue-PAST
“Enduring pain, (I) kept running.”

(2) Additive
zyoon-wa akaruku te kinben-da.
Joan-TOP be-cheerful-te diligent COPULA-PRES
“Joan is cheerful and diligent.”

(3) Temporal Sequence
gogo-wa tegami-wo kaite, ronbun-wo yonda.
afternoon-TOP letter-ACC write-te thesis-ACC read-PAST
“In the afternoon, (I) wrote letters and read the thesis.”

(4) Cause-Effect
taihuu-ga ki-te, i-e gahai-sareta.
typhoon-NOM come-te houses-NOM destroy-PASSIVE-PAST
“A typhoon came, and houses were destroyed.”

The examples are borrowed from (Hasegawa, 1996).
3 Organization of the Coherence Relations

Although the semantic relations between the te-linked constituents are diverse, not all relations implicated by parataxis can be expressed by te-linkage (Hasegawa, 1996). For example, if the clauses equivalent to I sat down and The door opened are presented paratactically in Japanese, the interpreter naturally reads in a Temporal Sequence relation, just as in English. But this relation is not an available interpretation when the clauses are linked by te. That is, among the relations potentially implicated by two copresent clauses, some are filtered out by te-linkage.

We presume that the inherent meaning of te is “togetherness.” The only relations that fit with this meaning are possible to arise within te-linkage. The notion of “togetherness” can be divided into two categories according to the temporal properties of relations. One in parallel and the other in series. In the former, two events occur simultaneously or two states hold at the same time, while in the latter, two events occur successively.

These two categories are further divided into smaller categories according to the event structures of conjuncts. The category of sequential relations contains both Cause-Effect and Temporal Sequence. When two events which are linked solely by temporal sequentiality are expressed via te-linkage, the conjuncts must share an agentive subject. Thus, causation and one person’s volitional acts are sufficient to be recognized as togetherness.

On the other hand, in order for the category of parallel occurrence of events to be compatible with te-linkage, they must be homogeneous in some sense. One such example is the case where a thing has two different properties (Additive) and another is the cases where two different things have similar properties or are engaged in similar events (Contrast). As for the Additive relation, the subject of the second conjunct is often omitted since it is the same as that of the first. In addition, both predicates of the conjuncts are static — adjectives or stative verbs — because they have no temporal boundaries as opposed to events and can easily hold at the same time within one person. As for the Contrast relation, the subjects of the conjuncts must be different from each other and hence both of them are explicitly mentioned (often marked with the contrastive wa). In general, the similarities of the predicates appear as the syntactic parallelism as the example (6) shows.

The other sub-category of the parallel occurrence of events is “accompaniment,” where the second clause is foregrounded and the first backgrounded. The prototypical instance of this category is the case where the first clause denotes some manner of causation, since the means can be interpreted as a kind of causation. This is exemplified by Japanese doosite (why/how) as follows:

(18a) densya-de (means)
   “by train”

(18b) aitakatta-kara (reason)
“since (I) want to meet (you)”

(18b) expresses the reason why the speaker came to
the hearer — “the wish to meet the hearer caused
him/her to come.” Thus, this relation associates the
two extremes i.e., parallelism and sequentiality.

Finally, the Concession is closely related to both
Cause and Contrast. In the Concession relation, the
first clause implies something and the second clause
denies it. The implied states or events are often those
to be caused by the events or states denoted by the
first clause, and then denied and contrast with the
second clause.

The whole organization is shown in Figure 1. Note

|STEP1| Assume part of semantic structures of the
      |conjuncts by reverse linking |
|STEP2| Unify them with a verb’s semantic structures |
|STEP3| Infer the most feasible relation between
      | them |

In STEP1, part of the semantic structure of each clause is abductively assumed by applying linking rules backward. The linking rules are regular ways of...
mapping open arguments — i.e., variables of semant- 
ic structures whose referents can be expressed syn-
tactically by a phrase within the same clause as the 
predicate — onto grammatical functions or under-
lying syntactic configurations by virtue of thematic 
roles (thematic roles are positions in a structured 
semantic representation). In the case of Japanese, 
they are triggered by case particles. In STEP2, 
the verb’s semantic structures are invoked and uni-
fied with the outputs of STEP1. The examples of 
the linking rules and verbs’ semantic structures are 
shown in Figure 2 and 3 respectively.

However, since the real texts contain far more 
complexity and ambiguity than the examples given 
in this paper, we have to correct the outputs of the 
processes manually (the gapped arguments are filled 
by hand). We now focus on the processes that cal-
culate the coherence relations.

4.2 The Properties Relevant to the 
Coherence Relations

What is essential for recognizing the coherence rela-
tion between clauses is that the constituents of one 
clause bear certain kind of structural relationship to 
those of the other. Although there are an infinite 
number of situations, there seems to be only a small 
number of properties relevant to the coherence rela-
tions that can hold between them. They are:

1) the identity and agentivity of the subjects in 
   the two clauses
2) the thematic and aspectual properties of the 
event denoted by each clause
3) canonical events associated with the noun that 
is relevant to both clauses

Before going through the use of these properties, 
let’s consider the other information which affects our 
construal of the relations.

There are some adverbials or fixed expressions 
which coerce the interpretation into the specific rela-
tion. In addition, there are narrow-range verb 
classes which specialize the implicated relation by 
virtue of their inherent meaning. For example,

verbs that take a temporal NP as the subject and 
means “the passage of time” such as sugiru(pass 
away), owaru(end), keikasuru(elapse), etc., imply 
the Temporal Sequence relation when followed by te. 
Verbs that express “using” such as tukau(use), siy-
ousuru(make use of), katuyousuru(apply), etc., im-
ply the Means-End relation. They are summarized 
in Table 1. In Table 1, [TE] means temporal ex-
pressions such as days, months, years, centuries, etc. 
The verbs and fixed expressions appear in the first 
clause, while the adverbials in the second. These 
fixed expressions should be listed as a unit in the 
lexicon.

When these expressions appear in the test sen-
tences, we can identify the relation regardless of the 
procedure described below. Otherwise, we have re-
course to the aforementioned properties.

4.3 The Prototypes and the Extensions

In the previous study, we have classified verbs into 30 
semantic categories, and for each category we have 
given a lexical conceptual structure (LCS) represen-
tation (Oishi and Matsumoto, 1997). Since the LCS 
representation involves lexical decomposition (Jack-
endoff, 1990), we can utilize the verb internal seman-
tic structure so as to calculate coherence relations in 
a fairly principled way.

As mentioned in the introduction, we consider 
each relation as a category. Categories cannot be 
defined in terms of necessary and sufficient condi-
tions, but rather each instance is categorized according 
to its similarity to the prototypes of the cate-
gories (Rosch, 1973; Lakoff, 1987; Taylor, 1989).

We define a prototypical structure for each rela-
tion by means of the predicates used in the LCSs as 
follows:

- **Circumstance**
  \[x \text{ACT}_2 \text{WITH} [x \text{BE} z_1] \]

- **Additive**
  \[[x \text{BE} z_1] \text{AND} [x \text{BE} z_2]_2 \]

- **Temporal Sequence**
  \[[x \text{GO TO} z_1] \text{THEN} [x \text{GO} \text{FROM} z_1 \text{TO} z_2]_2 \]

<table>
<thead>
<tr>
<th>relations</th>
<th>categories</th>
<th>examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporal Sequence</td>
<td>passage verbs</td>
<td>sugiru(pass away), keikasuru(elapse) ...</td>
</tr>
<tr>
<td></td>
<td>ending verbs</td>
<td>owaru(end), oeru(finish) ...</td>
</tr>
<tr>
<td></td>
<td>continuing verbs</td>
<td>tuzuku(continue), hikituzuku(follow) ...</td>
</tr>
<tr>
<td></td>
<td>adverbials</td>
<td>sonogo(after that), imadeha(nowadays) ...</td>
</tr>
<tr>
<td></td>
<td>fixed expressions</td>
<td>[TE]ni-natte(set in), [TE]hodo-site(after) ...</td>
</tr>
<tr>
<td>Means-End</td>
<td>using verbs</td>
<td>tukau(use), siyousuru(make use of) ...</td>
</tr>
<tr>
<td></td>
<td>fixed expressions</td>
<td>ni-yotte(by means of) ...</td>
</tr>
<tr>
<td>Cause-Effect</td>
<td>fixed expressions</td>
<td>dake-atte(on account of), wo-ukete(given) ...</td>
</tr>
<tr>
<td>Circumstance</td>
<td>static relation verbs</td>
<td>sou(be parallel to), motozuku(be based) ...</td>
</tr>
</tbody>
</table>
• Cause-Effect
\[[x \text{ ACT ON } y_1] \text{ CAUSE } [y \text{ BECOME } z_2]\]

• Means-End
\[[x \text{ ACT}]_2 \text{ BY } [x \text{ ACT}]_1\]

• Contrast
\[[x \text{ ACT}]_1 \text{ WHILE } [y \text{ ACT}]_2\]

• Concession
\[[x \text{ ACT ON } y_1], \text{ BUT } [y \text{ NOT BECOME } z_2]\]

Here, WITH, AND, THEN, etc., are mnemonic names for the relations and each can be considered as a function that takes two events or states as its arguments and returns a coherent event or state. We use the infix notation for each function rather than prefix. The square brackets identify the surface ordering of the clauses linked by a mantic structure of a clause and their subscripts denote the temporal ordering of the clauses linked by te. ACT, BE, GO, and BECOME are also functions and they correspond to actions, states, movement, and inchoatives respectively. They express broad-range classes of the events which are constructed by the previous steps (see Figure 3). The whole structures incorporate the identity between the subjects of two clauses by the variables x and y. Agentivity of each subject is implied by the types of the events: ACT > GO > BECOME > BE.

Often, these prototypical structures are lexicalized and expressed by a single clause. For example, the Cause-Effect relation is lexicalized into accomplishment verbs (Talmy, 1985) and the Means-End relation can be expressed by an adjunct event noun followed by the case particle de. They must be extended so that they can cover wider range of instances of te-linkage. The result of the extension is shown in Table 2 (for cases each of which shares a subject) and Table 3 (for cases each of which has distinct subjects), where each column corresponds to the type of the event in the first clause and each row to the second. The prototypes are boldfaced and they are extended to the other boxes with some directions and constraints.

For example, the Temporal Sequence relation has a prototype structure, which is roughly read as “someone goes to somewhere, and then he/she goes (from there) to elsewhere.” This expresses our common sense that one person cannot move along two different paths at the same time, which implies that the two movements by a person must be sequential. This prototype is extended so as to cover such situations as “someone goes to somewhere, and then he/she does something/becomes something/stays there” or “someone does something/become something/stays somewhere, and then he/she goes to elsewhere.” They are expressed by vertical and horizontal extensions of the prototype in Table 2. The movements involved in these situations are local and the other events must be done volitionally by the same person. Another extension covers situations where “someone does something, and then he/she does something else.” This is based on the fact that one person cannot generally engage in two actions at the same time. Of course, any type of events may occur sequentially. However, there exists the constraint on the fitness with te-linkage as mentioned in the previous section.

The explanation for the other relations is detailed in (Oishi, 1998).

As a result of the extensions, many boxes have two or more relations. Notice that the nearer relations in the organization tend to be in the same boxes. To discriminate among them, we specify for each combination of event types such algorithm as follows (below, I(i,j) means that two clauses share an subject and D(i,j) means that two clauses have distinct subjects, where i is the event type of the first clause and j the second):

• I(ACT,ACTION), I(ACT,GO)
  If either clause contains the expressions which fix the temporal boundary, then Temporal Sequence;
  else if the verb of the first clause involves a manner component, then Circumstance;
  otherwise, Means-End.

• I(ACT,BECOME)
  If the second event is psychological, then Cause-Effect;

Table 2: The combinations of event types (identical subjects)

<table>
<thead>
<tr>
<th>2nd clause</th>
<th>ACT</th>
<th>GO</th>
<th>BECOME</th>
<th>BE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACT</td>
<td>Means</td>
<td>TempSeq</td>
<td>TempSeq</td>
<td>Circum</td>
</tr>
<tr>
<td>GO</td>
<td>TempSeq</td>
<td>TempSeq</td>
<td>TempSeq</td>
<td>TempSeq</td>
</tr>
<tr>
<td>BECOME</td>
<td>Cause</td>
<td>TempSeq</td>
<td>Cause</td>
<td>Circum</td>
</tr>
<tr>
<td>BE</td>
<td>Cir(manner)</td>
<td>TempSeq</td>
<td>Additive</td>
<td>Circum</td>
</tr>
</tbody>
</table>

Table 3: The combinations of event types (distinct subjects)

<table>
<thead>
<tr>
<th>2nd clause</th>
<th>ACT</th>
<th>GO</th>
<th>BECOME</th>
<th>BE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACT</td>
<td>Contrast</td>
<td>Circum</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO</td>
<td>Contrast</td>
<td>Circum</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BECOME</td>
<td>Cause</td>
<td>Contrast</td>
<td>Cause</td>
<td>Circum</td>
</tr>
<tr>
<td>BE</td>
<td>Concession</td>
<td>Contrast</td>
<td>Circum</td>
<td></td>
</tr>
</tbody>
</table>
I(GO,BECOME)
If the second event is psychological, then Cause-Effect;
otherwise, Temporal Sequence.

I(BECOME,GO)
If the first event is perceptual, then Cause-Effect;
otherwise, Temporal Sequence.

I(BE,GO)
If either clause contains the expressions which fix the temporal boundary, then Temporal Sequence;
otherwise, Circumstance.

I(BE,BECOME)
If the second event is psychological, then Cause-Effect;
otherwise, Circumstance.

I(BE,BE)
If the second state is psychological, then Cause-Effect;
else if the both predicates are property-denoting adjectives or nouns, then Additive;
otherwise, Circumstance.

D(BECOME,BECOME)
If the both subjects are marked with wa, then Contrast;
otherwise, Cause-Effect.

D(BE,BE)
If the both subjects are marked with wa, then Contrast;
otherwise, Circumstance.

On the other hand, there remain some boxes blank. They should be resolved by using the third property — the canonical events associated with the noun that is relevant to both clauses. The generative lexicon will serve the purpose (Pustejovsky, 1995). At present, however, we have not yet fully implemented the lexicon for nouns. Therefore, we give the Circumstance relation as a default.

5 Experiment and Discussion
An experiment of recognizing coherence relations of te-linkage were done for 280 sentences which were randomly extracted from EDR Corpus (EDR, 1995). The analysis results are shown in Table 4, where the coherence relations in the sentences were classified into 7 categories by authors and compared with the outputs of the program.

The relations are not balanced in number. This seems to be due to the genre of texts from which the test sentences were picked up (most of them were news articles). The numbers in parentheses show those of test sentences that matched with the fixed expressions in Table 1.

The precision on the whole is 82%. This shows that to a large extent we can cope with the problem to recognize the coherence relations between clauses (at least when linked by te), given the event types of the clauses and the fixed expressions in the lexicon.

Most of errors are caused by ambiguity of the relation. There were many examples which were difficult even for humans to make clear judgements. This reflects the fact that the coherence relations do not have definite borders.

However, there were some errors which show a crucial limitation of our method. This appears as the bad marks in both precision and recall for the Concession relation, even though the number is small. For example, there is a test sentence such as follows:

(19) ano hito-wa 82sai-ni natete, annani koukisin ippai-da.
that person-TOP 82-years-old-DAT become-te,
so curiosity be-full-PRES
"Although that person is 82 years old, (he/she) is full of curiosity."

Table 4: The results of the experiment

<table>
<thead>
<tr>
<th>coherence relations</th>
<th>judgement by human(a)</th>
<th>output of program(b)</th>
<th>number of agreements(c)</th>
<th>recall(%) c/a×100</th>
<th>precision(%) c/b×100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporal Sequence</td>
<td>89</td>
<td>81(46)</td>
<td>79</td>
<td>89</td>
<td>98</td>
</tr>
<tr>
<td>Circumstance</td>
<td>75</td>
<td>83(22)</td>
<td>63</td>
<td>84</td>
<td>76</td>
</tr>
<tr>
<td>Cause-Effect</td>
<td>64</td>
<td>58(13)</td>
<td>48</td>
<td>75</td>
<td>82</td>
</tr>
<tr>
<td>Means-End</td>
<td>45</td>
<td>48(12)</td>
<td>34</td>
<td>76</td>
<td>71</td>
</tr>
<tr>
<td>Additive</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Concession</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>33</td>
<td>20</td>
</tr>
<tr>
<td>Contrast</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>Total</td>
<td>280</td>
<td>280(92)</td>
<td>229</td>
<td>82</td>
<td>82</td>
</tr>
</tbody>
</table>
Since the combination of the event type here is I(BECOME,BE), our program gave it the Circum-
stance relation as a default. However, we know that in general the person who is 82 years old is not so curious, therefore the Concession relation arises. Thus, our common sense knowledge is crucial to our recognition of the coherence relations. In (Hovy and Maier, 1993), they classified the Concession relation as interpersonal (i.e., author-and/or addressee-related) rather than ideational (i.e., semantic), since they defined it as “one of the text segments raises expectations which are contradicted/violated by the other.” The use of interpersonal relations is predicated mainly on the interests, beliefs, and attitudes of addressee and/or author. To deal with this problem, we must incorporate the notion of intentional structure and focus space structure (Grosz and Sidner, 1986).

Since we have focused on te-linkage in this paper, we need not to consider how clauses are combined. However, to detect the discourse structure, we need to extend the method so as to deal with the relations between sentences. We must estimate some kind of reliable scores among possible segments and choose the relation having the maximum score (Kurohashi and Nagao, 1994). These issues remain to be studied in the future.

6 Summary

Since the semantic relations exhibited by te-linkage vary so diversely, it has been claimed that the interpreter must infer the intended relationship on the basis of extralinguistic knowledge. The particulars of individual common sense knowledge are crucial to understanding any discourse (Hobbs et al., 1993; Asher and Lascarides, 1995). Nevertheless, one can, through the use of the relevant structures of events, eliminate a very large number of rules for calculating the plausible relations.

Although we have concentrated on te-linkage in this paper, we consider that the method can be applied to pure parataxis with necessary modifications. For the relations we have examined are not attributable to the meaning of te itself (though it restricts the range of them), but are implicated by the linked conjuncts. The same is true of English and. In both and- and te-linkage, the perceived coherence relations are present even if the linked constituents are in pure parataxis without and or te. Thus, this approach can be extended so as to detect the whole discourse structure, though further study must be done to examine all relations.

References

N. Asher and A. Lascarides. 1995. Lexical disambigu-


