ENGLISH-STYLE AND CHINESE-STYLE TOPIC: A UNIFORM SEMANTIC ANALYSIS*

Hui Xu

Department of Computational Linguistics
University of the Saarland, 66041 Saarbrücken, Germany
xu@mail.coli.uni-sb.de

ABSTRACT

This paper provides a uniform semantic framework for Topic-Comment Structures in Chinese. It is assumed that the topic element relates to an anaphor occuring in the comment. The comment which contains an unresolved anaphor is semantically underspecified. The semantic interpretation of topic-comment constructions results from the resolution of the anaphoric relation between topic and comment. In the case where there is no explicit anaphor in the comment, a bridging inference takes place. The proposed analysis is modelled in a DRT-based framework. The introduction of a DRS segmented for topic-comment articulation, called TC-DRS, serves to compute the topic-comment relation within the discourse semantic theory DRT compositionally.

1 INTRODUCTION

Topic-comment structures (TCS) have attracted much research in a wide range of languages. By studying Chinese data, the difference between the topic "English-style", which is co-referential with some (overt or covert) element in the comment, and the topic "Chinese-style", which is related to the comment by a relation of aboutness, is quite striking.[1] Both types of topic are exemplified in (1) and (2) respectively.

(1) *[Zhe ben xiaoshuo]₁ Zhangsan kanwan le e₁.*
    this CL novel   Zhangsan read finish Perf
    (This novel, Zhangsan has finished.)

(2) *Neike shu, yezi da.*
    that tree leaves big
    (That tree, the leaves are big.)

In (1) there is an empty trace, interpreted as a zero anaphor that is bound by the topic. By contrast, there is no such syntactic position inside the comment associated with the topic in (2). Instead, topic and comment are related with each other pragmatically. Therefore, the topic in (2) is considered by some scholars to be "base-generated" in front of the sentence, and to have an "aboutness"-relation to the remaining sentence (e.g. [5]). However, a definition of "aboutness" is still not given. A pure syntactic definition of topic as an extraposition fails in the face of examples

---

[1]This dichotomy originated with [3]. Chinese-style topic constructions are also widespread in other topic-prominent languages such as Japanese, Korean, Lahu and Lisu ([5]).

such as (2) while the relational approach advocated in [10] leaves open how to characterize the relation between topic and comment in (2). Furthermore, the semantics of the Chinese-style topic construction remains unexplored.

In this paper, we propose a uniform semantic framework for interpreting the TCS in Chinese. We assume that the topic constituents bind anaphors inside the comments, such as pro-forms and implicit anaphora. Because the comments contain unresolved anaphors, they are semantically incomplete or underspecified. The semantic interpretation of TCS is given by anaphora resolution. In the absence of a direct anaphoric relation as in (2), a bridging inference is triggered. The proposed interpretation will be modelled in a discourse semantic framework by extending Discourse Representation Theory (DRT) with topic-comment articulations. We will show that this approach can also capture multiple-topic constructions and topic chains.

The paper is structured as follows. Section 2 gives a brief review of related approaches. It shows that to capture the topic-comment relations we have to take a closer look at the semantic properties of the TCS. Section 3 discusses how bridging works to account for the semantical relations in the TCS. Section 4 presents a DRT-based framework of topic-comment structure, which is applied to some examples of linking and bridging as well as to some dangling topic constructions involving ellipsis in section 5. We extend our framework to topic chains in section 6 and to multiple-topic constructions in section 7. Section 8 concludes with pointers for future research.

2 A BRIEF REVIEW

In this section we start by discussing the syntactic approach of [8] which considers topicalized sentences as derivation from a more basic structure. We show that such a reanalysis cannot capture the great variety of topic-comment relations. Furthermore we argue that the relational approach of [10] which relate the pragmatic function and the syntactic properties of topic fails to specify the semantic relationship of the Chinese-style topic to the comment.

The Derivation Approach of [8] considers the TCS as a grammatical device invoked to fulfill certain discourse functions. Shi's claim is that all topics are related to a structural position inside the comment. The comment itself is thus an open clause with a gap, namely, a position filled by an empty category or a resumptive pronoun, which is co-referential with the topic. That is, he attempts to define the concept of aboutness syntactically. The major supporting argument of [8] is that subject and object are not the only positions which the topic can relate to. The oblique object of adjunct adverbial PPs for example, often serves as the position bound by the topic as illustrated in (3). (3) can be seen as derived from (4).

(3) *Zhe   jian   shiqing   ni   bu   neng   guang   mafan   yi   ge   ren.*
     this   CL   matter   you   not   can   only   bother one   CL   person
     (For this matter, you cannot only bother one person.)

(4) *Ni   wei   zhe   jian   shiqing,   bu   neng   guang   mafan   yi   ge   ren.*
     you   for   this   CL   matter   not   can   only   bother one   CL   person
     (You cannot only bother one person for this matter.)

The derivation approach attempts to show that Chinese-style topic constructions are derived from some basic sentence structure via syntactic operations like movement and deletion. However, such a claim is not always supported by the data. As (5) shows, there is not always a more basic sentence from which to derive the TCS. Although (5) can be paraphrased as (6), (6) is not to be considered as the basic form that (5) is derived from, because (5) sounds more natural than (6).

(5) *Hua,   Mali   zui   xihuan   meiguihua.*
     flower Mary   most like   rose
     (Concerning flower, Mary likes roses the most.)

(6) *Suoyou de hua zhong Mali zui xihuan meiguihua.*
all   MOD   flower   among   Mary   most   like   rose
(Among all flowers, Mary likes roses the most.)

Moreover, in the interpretation of (7), it is the teacher Wang who cannot come. By contrast, (8) is ambiguous as either teacher Wang or his daughter can be understood as not being able to come. The "re-interpretation" approach fails to account for this difference.

(7) *Wanglaoshi$_1$, nuer$_2$ bing le, e$_1$ bu neng lai.*
teacher Wang   daughter   ill   ASP   not   can   come
(As for teacher Wang, his daughter is ill, (he) cannot come.)

(8) *Wanglaoshi$_1$ de nuer$_2$ bing le, e$_1$/e$_2$ bu neng lai.*
teacher Wang   MOD   daughter   ill   ASP   not   can   come
(The daughter of teacher Wang is ill, (he/she) cannot come.)

The Relational Approach of [10] generalizes the TCS with gap (1) and without gap (2), and define the topic as the element X in

$$[_{S'} X [_S ... Y ...]],$$ where X is a major category and Y, possibly empty, is related to X.

The related constituent (empty category or lexical item) in the comment is interpreted as a pro-form whose antecedent is the topic. (1) is then interpreted as *as for this novel, Zhangsan has finished it*. This interpretation obviously corresponds to our intuition. However, [10] do not explore how to apply the topic definition to (2). In (2) there is a non-empty lexical item relating to the topic, but in an indirect way. The present approach originates from the underlying assumption of [10]. We give a more precise semantic analysis of TCS and propose that the semantics of TCS involve an anaphoric relation which must be resolved during comprehension.

3   TCS, BRIDGING, AND ANAPHORA

The relations between topic and comment, which are pragmatically defined as "aboutness", appear semantically in great variety. In (2) the leaves are part of the tree introduced as topic. In (9) the topic gives the place where the speaker grew something. In (10) the female students whom Peter gave books to belong to the students mentioned in the topic.

(9) *Nei ge huayuan, women yiging zhongshang e le.*
that   CL   garden   we   already   grow   ASP
(That garden, we already grew.)

(10) *Zhexie xuesheng, Bide gei le meige nusheng yi ben shu.*
these   students   Peter   give   ASP   every   female students   a   CL   book
(These students, Peter gave every female students a book.)

Nevertheless, these examples have in common that each topic has a bridging relation to an element inside the comment. These bridging relations are an essential part of the sentential content in the sense that without this information, the lack of connection between topic and comment would make the sentences incoherent and therefore incomprehensible. To capture the semantic of TCS, we have to build a bridge between topic and comment.

Bridging is defined in [1] as "an inference that two objects or events that are introduced in a text are related in a particular way that isn't explicitly stated." The most investigated bridging relation is triggered by functional anaphora which appear in a context where no actual antecedents are present, but the network of discourse referents that are previously established provides the right

antecedent. In TCS, however, bridging originally arises out of a coherence constraint on the TCS, as it checks to see if a comment can be attached to a topic preceding it. Semantically this means that there is an anaphor inside the comment which is anaphoric to the topic. Precisely, the topic gives an antecedent which should be picked up in the comment. The English-style topics have a direct anaphoric relation to the comment while the Chinese-style topics have an indirect reference to the comment. Bridging occurs if no explicit anaphor is present inside the comment.

We adopt the view of [2] who consider bridging as a possibility for anaphora resolution. Anaphora resolution can take place in three different ways: linking to an accessible[2] suitable antecedent in discourse, bridging to an accessible antecedent with accommodated material or accommodating the anaphoric information into the discourse. Linking is prefered to bridging, and bridging to accommodation. To handle functional anaphora, [2] develop a theory of bridging by extending [7]'s work with lexical knowledge. The strategy is to include more lexical information represented as qualia structure ([6]) in the discourse, so that functional anaphora can link to objects of this additional information. Qualia information is normally not accessible and does not affect the truth-conditions of discourse. If necessary, it is put forward to the surface by a process called coercive accommodation, and is then available for anaphora resolution.

Lexical knowledge is also utilised to meet the coherence constraints imposed by the TCS. As mentioned above inside TCS, bridging can be triggered by the coherence constraint, as this requires that some element u of the comment relates in some way B to the topic. B can be linking or bridging. In the second case, the lexical information about u is involved in the interpretation. (2), for example, can be interpreted as a comment to the topic *that tree* if one can infer that (i) leaves are part of some bigger plant, (ii) the leaves which are big are part of x, and (iii) x is identical with *that tree*. (i) is implied in lexical knowledge of leaves. (ii) results from accommodating (i) into the comment. (iii) is the anaphora resolution of x. In section 4, we show how the semantics of TCS can be computed in a discourse semantic framework.

4   DISCOURSE SEMANTICS OF TCS

In this section we present a semantic analysis of TCS in a version of DRT. In TCS, a topic constituent outscopes the whole sentence or a sequence of sentences. It is however anaphorically related by anaphora to the subsequent sentences. So the semantic representation of TCS must give an account of this semantic effect. The representational framework DRT seems to be an appropriate formalism because the discourse information of discourse referents is processed in a dynamic way. To interpret the TCS, we extend DRT with topic-comment articulations. After a brief description of how DRT works and how anaphora is handled in DRT, we show how a DRT approach can be used to capture the TCS.

4.1   Discourse Representation Structures

In DRT sentential meaning is interpreted with respect to the context in which the sentence is embedded. The discourse information is represented in the form of a discourse representation structure (DRS) built algorithmically on the basis of the syntactic analysis. A DRS K consists of a set of discourse referents (DR) $U_K = \{x, y, z, ..., X, Y, Z, ...\}$[3] and a set of conditions about these DRs C_K. DRS-Construction processes in a dynamic way in that the current sentence is added to the DRS constructed for the previous discourse thus yielding a new DRS as a context for the interpretation of subsequent sentences. To represent anaphoric information, a new DR is introduced. As condition this DR has to be linked with some accessible DR. [2] introduce a special type of DRSs called α-DRSs to mark anaphoric information. The α-operator quantifies over a DR which should be bound by a DR established before. DRSs that contain α-DRSs are unresolved DRSs. α-DRS resolution results from linking the α-DR to an accessible and suitable DR in the

---

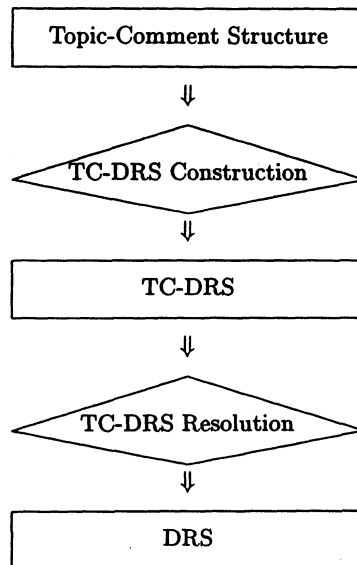[2]Accessibility is defined in [4].

[3]Upper letters represent plural DRs.

DRS of the previous discourse, or from building a bridge between the α-DRS and an accessible DR in the previous DRS by accommodating related lexical information.

DRSs with no unresolved α-DRSs are proper DRSs. Proper DRSs can be interpreted in a certain model: they are true with respect to that model if they can be embedded in it ([4]). In this paper we will not discuss the model theory in more detail. Rather we will concentrate on the DRS construction and resolution process.

4.2   DRS-Construction of TCS

The DRS-construction of TCS is implemented in two steps: firstly, topic and comment are translated into a TC-DRS. Secondly, the TC-DRS is resolved, yielding an anaphoric relation with the topic discourse referents (T-DR), i.e. the discourse referent for the topic contituent, as antecedent. The algorithm for DRS-construction of TCS can be illustrated in the following diagram.

```
┌─────────────────────────────┐
│   Topic-Comment Structure   │
└─────────────────────────────┘
               ⇓
        <  TC-DRS Construction  >
               ⇓
┌─────────────────────────────┐
│           TC-DRS            │
└─────────────────────────────┘
               ⇓
        <  TC-DRS Resolution  >
               ⇓
┌─────────────────────────────┐
│            DRS             │
└─────────────────────────────┘
```

A TC-DRS is a DRS segmented for topic-comment articulation. It is defined as a DRS with a separate part for T-DRS and C-DRS, i.e. temporary DRSs[4] representing the topic and comment information respectively. T-DRS and C-DRS can only occur in a TC-DRS simultaneously. The vocabulary of TC-DRS is the same as defined in standard DRS ([4]) except that T-DR are written as variables with a t-superscript, e.g. x^t, y^t. In what follows, we define the syntax of T-DRS, C-DRS and TC-DRS.

> Definition: **Syntax of Topic DRS (T-DRS)**
> A DRS $K = <U, C>$ is a T-DRS, iff U contains exactly one topic discourse referent, or C contains an α-DRS whose α-DR is a topic discourse referent.
>
> Definition: **Syntax of Comment DRS (C-DRS)**
> K is a C-DRS, iff K is a simple DRS or a TC-DRS.
>
> Definition: **Syntax of Topic-Comment segmented DRS (TC-DRS)**
> $K = <U, C, K^T, K^C>$ is a TC-DRS, iff U is a set of discourse referents, C is a set of DRS-conditions, K^T is a T-DRS and K^C is a C-DRS.

4.2.1   *TC-DRS Construction*

The construction of T-DRS and C-DRS is according to [4] and [2], except for the t-marking of T-DR. Assuming that the syntactic tree of a TCS S' is

---

[4]Temporary DRSs are represented by dotted boxes. Proper DRSs may not contain any temporary DRSs as these are not assigned any semantics and should be resolved by discourse information.

$$\begin{array}{c} \text{S'} \\ \diagup \diagdown \\ \text{TOPIC} \qquad S \end{array}$$
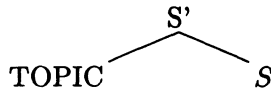
then the translation rule of TCS into a TC-DRS is formulated as below:

Translation Rule for TC-DRS:

  a. Translate *TOPIC* into a T-DRS K^T, where the DR introduced by the topic is a T-DR.

  b. Translate S into a C-DRS K^C.

  c. The current DRS $K = < U, C >$ is updated to a TC-DRS $K' = < U, C, K^T, K^C >$.

4.2.2 *TC-DRS Resolution*

By the resolution of TC-DRS, two cases are to be distinguished: in the first case the C-DRS contains an explicit anaphor (represented as α-DRS) relating to the T-DR and a direct linking is allowed; in the second case, the anaphor relating to the T-DR should be accommodated by bridging inference. If it succeeds in one of these ways, a normal DRS results. Otherwise, an error message is returned. The resolution of TC-DRS is a partial function from TC-DRS to DRS.
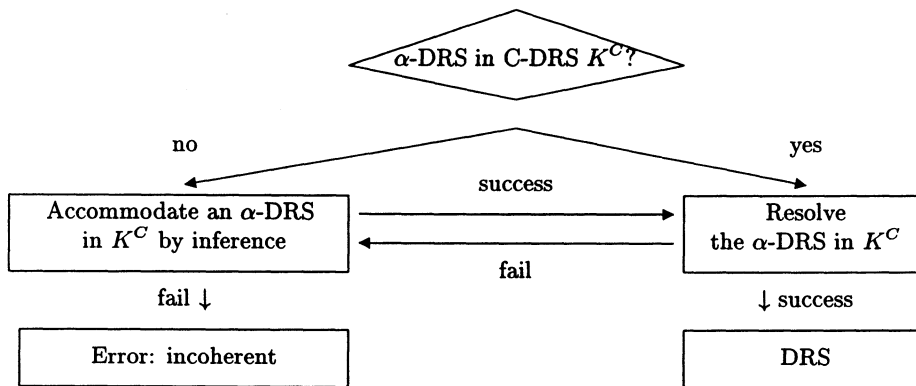
Resolution Rule for TC-DRS:
Let $K^{TC} = < U_K, C_K, K^T, K^C >$ be a TC-DRS where $K^T = < U_{K^T}, C_{K^T} >$, $K^C = < U_{K^C}, C_{K^C} >$. If there is an embedded α-DRS $\alpha : K_\alpha$ in C_{K^C}, go to ii); else go to i)

  i. Accommodate an α-DRS $\alpha : K_\alpha$ into K^C according to the bridging inference rule. Then go to ii).

  ii. Check whether the T-DR x^t in U_{K^T} is a suitable antecedent for the α-DR y in C_{K^C}. If this is the case, K^C is updated to $< U_{K^{C'}}, C_{K^{C'}} >$ where $U_{K^{C'}} = U_{K^C} \cup U_{K_\alpha}$ and $C_{K^{C'}} = C_{K^C} \cup C_{K_\alpha} \cup \{y = x\} - \alpha : K_\alpha$. Then go to iii). Otherwise, go to i).

  iii. The TC-DRS K^{TC} is updated to a DRS $K' = < U_K \cup U_{K^T} \cup U_{K^{C'}}, C_K \cup C_{K^T} \cup C_{K^{C'}} >$. Substitute the T-DR x^t with a normal DR x without t-Marking.

To get the anaphoric relation between topic and comment, lexical information as well as world knowledge are involved. The bridging inference rule is formulated as below:

Bridging Inference Rule
If a relation $R(y, z)$ is inferable by lexical information or world knowledge of the C-DRS K^C, and T-DR x^t is a suitable antecedent for y or z, then accommodate $R(y, z)$ and an α-DRS for y or z respectively as conditions into K^C.

5 APPLICATION TO EXAMPLES

We now examine the examples discussed in the introduction in more detail. Firstly, we concentrate on the linking relation involved in explicit anaphors. Secondly, we discuss bridging inferences. After that, we give examples of topic-comment relations resulting from domain restriction of quantification. In addition, we briefly take up TCS involving other constructions.

5.1 TCS with Explicit Anaphoric Relation

Explicit anaphoric relations between topic and comment can involve pronouns or definite descriptions. The syntactic tree of (1) is shown in (11a). It is translated into TC-DRS in two steps: firstly, $TOPIC$ is translated into a T-DRS K^T and then S into a C-DRS K^C (11b). The empty pronoun in the object position introduces an α-DRS. Because the T-DR x^t is a suitable antecedent for the α-DR z, we get the resolved DRS (11c).

(11) a.
b.

c.
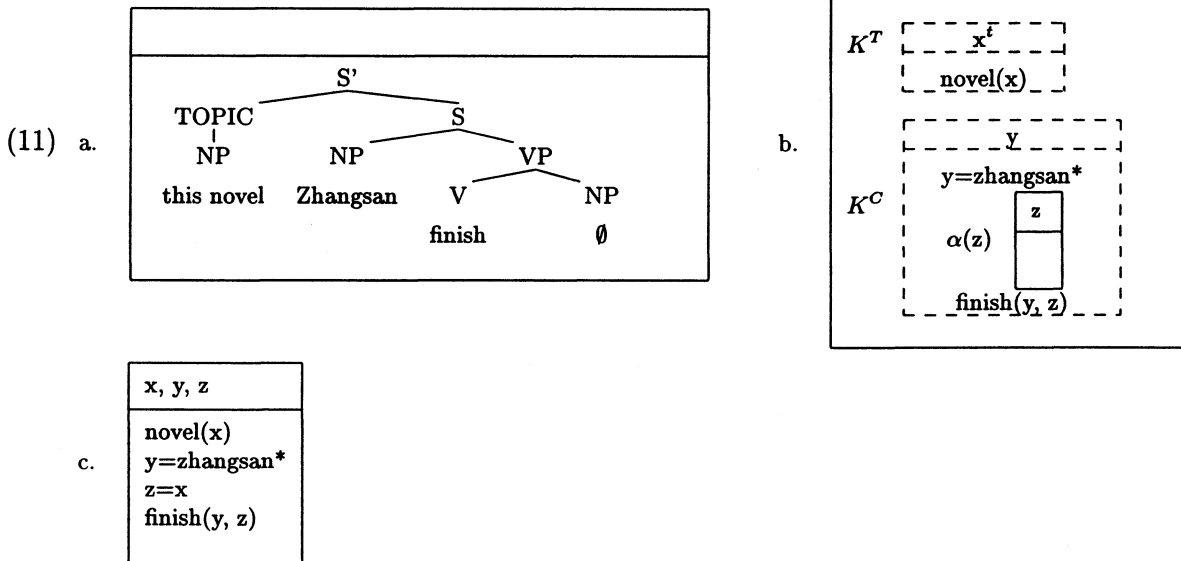
We assume that definite descriptions are anaphoric. In (12), the definite NP *this student* presupposes that there is a DR available in discourse which it anaphoricaly relates to. Linking the α-DR z to the T-DR x^t, we get the DRS (13b) interpreted as *Zhangsan knows a student who is Johan.*

(12) *Yuehan,   Zhangsan   renshi   zhege   xuesheng.*
      Johan   Zhangsan   know   this   student
      (Johan, Zhangsan knows this student.)

(13) a.
b.
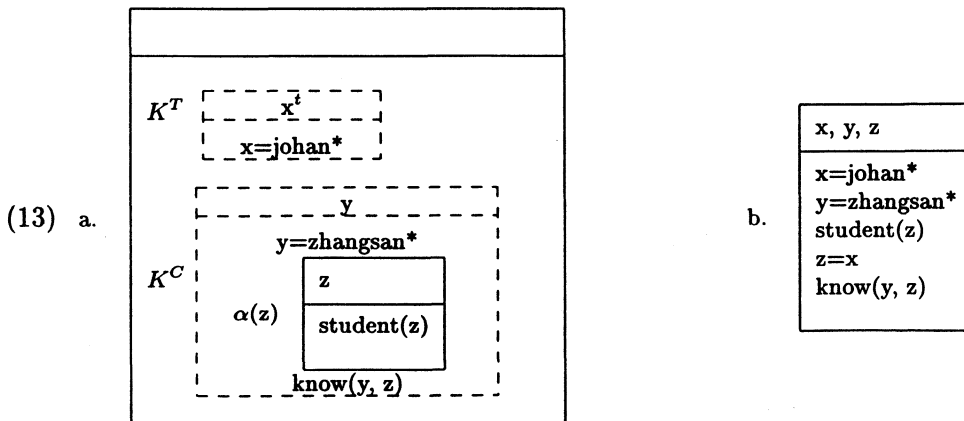
5.2 TCS with Implicit Anaphoric Relation

(2) is an example for bridging. As shown in (14a), there is neither a pronoun nor a definite description which introduces an α-DRS in K^C. So bridging is necessery to justify the TCS. Specifically, the background knowledge that *leaves are part of some plant, and tree is a plant* supports the bridge. (14a) is then updated to (14b). (14c) results from TC-DRS resolution.

(14) a.

```
        ┌ ─ ─ ─ ─ ─ ┐
    K^T │    x^t    │
        │  tree(x)  │
        └ ─ ─ ─ ─ ─ ┘
        ┌ ─ ─ ─ ─ ─ ┐
    K^C │     y     │
        │ leaves(y) │
        │  big(y)   │
        └ ─ ─ ─ ─ ─ ┘
```

b.

```
        ┌ ─ ─ ─ ─ ─ ─ ─ ┐
    K^T │      x^t      │
        │    tree(x)    │
        └ ─ ─ ─ ─ ─ ─ ─ ┘
        ┌ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ┐
        │          y            │
        │       leaves(y)       │
        │        big(y)         │
    K^C │      part_of(y, z)    │
        │               ┌─────┐ │
        │               │  z  │ │
        │        α(z)   ├─────┤ │
        │               │plant(z)│
        │               └─────┘ │
        └ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ┘
```
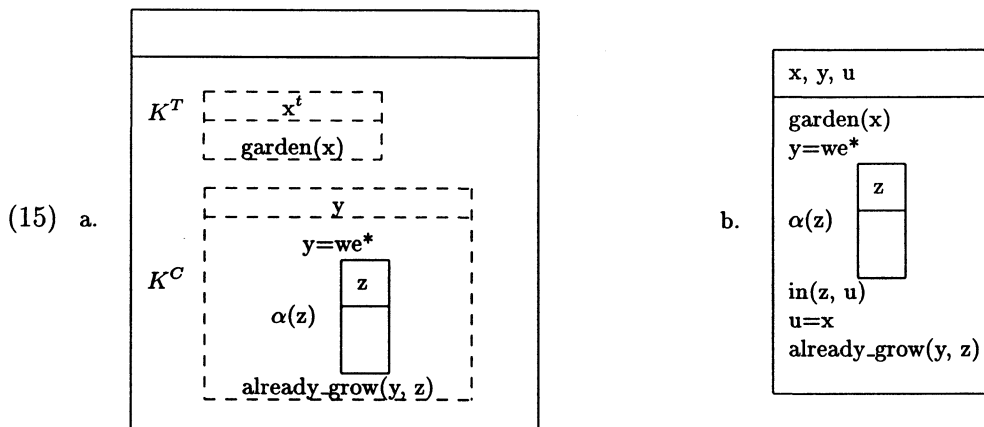
c.

```
┌──────────────┐
│   x, y, z    │
├──────────────┤
│ tree(x)      │
│ leaves(y)    │
│ big(y)       │
│ plant(z)     │
│ z=x          │
│ part_of(y, z)│
└──────────────┘
```
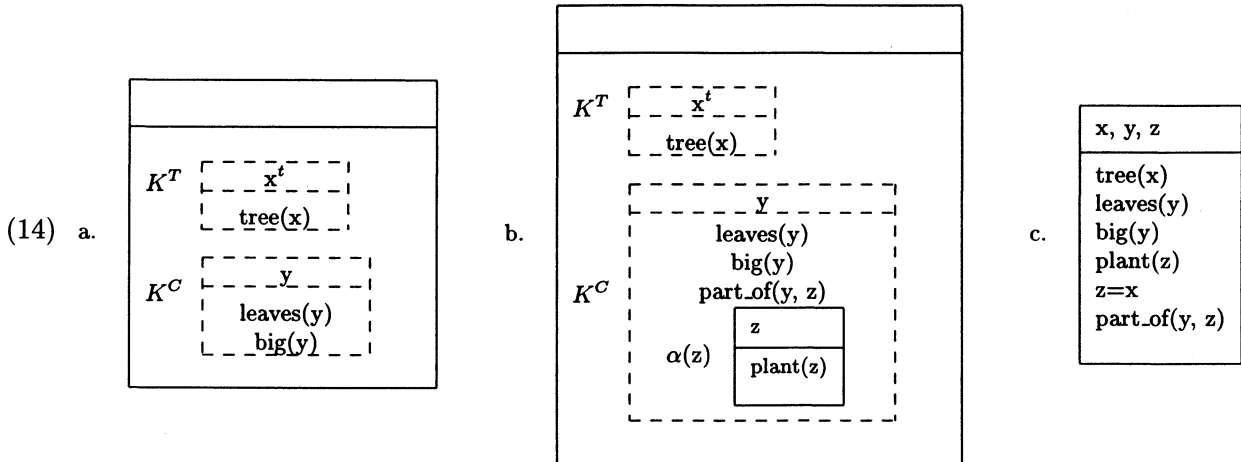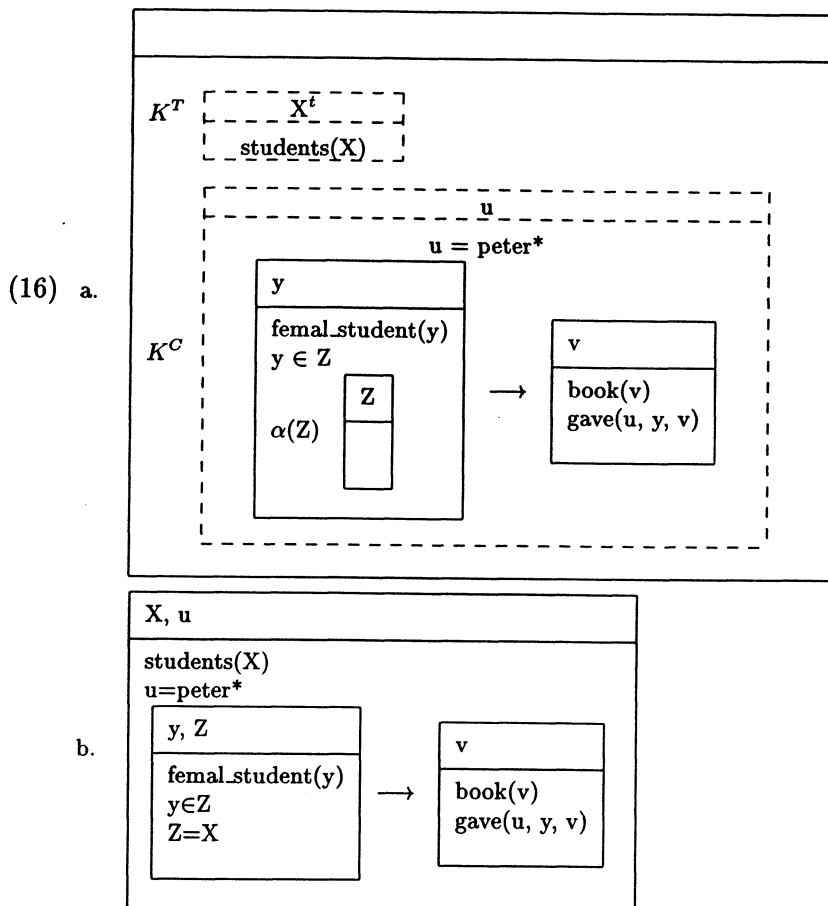
(9) is an other example of such dangling topics. As shown in (15a), the comment clause contains an empty pronoun, but it is not coindexed with the topic. In the natural interpretation we grew some plants in the garden, rather than the garden itself. The bridging is given by the background knowledge that plants are grown in a field such as a garden. (9) means that *that garden, we already grew something in it.* So the empty pronoun is interpreted as an anaphoric element relating to the topic by bridging, as shown in (15b).

(15) a.

```
        ┌ ─ ─ ─ ─ ─ ─ ─ ─ ┐
    K^T │       x^t       │
        │    garden(x)    │
        └ ─ ─ ─ ─ ─ ─ ─ ─ ┘
        ┌ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ┐
        │          y          │
        │        y=we*        │
        │             ┌─────┐ │
    K^C │             │  z  │ │
        │      α(z)   ├─────┤ │
        │             │     │ │
        │             └─────┘ │
        │  already_grow(y, z) │
        └ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ┘
```

b.

```
┌──────────────────────┐
│       x, y, u        │
├──────────────────────┤
│ garden(x)            │
│ y=we*                │
│             ┌─────┐  │
│             │  z  │  │
│      α(z)   ├─────┤  │
│             │     │  │
│             └─────┘  │
│ in(z, u)             │
│ u=x                  │
│ already_grow(y, z)   │
└──────────────────────┘
```

5.3 Domain of Quantification

Besides explicit anaphoric relations involving pronouns and definite descriptions, and implicit anaphoric relations provided by bridging, there is another kind of topic-comment relations given by quantificational domains. It is generally assumed that the domain of quantification depends on the context ([9]). It can be seen as an implicit argument of a quantifier since it is usually not uttered literally, but will be instantiated by the context information. In TCS, the domain of quantification is activated to be bound by the topic. In (10), the domain of the all-quantifier *every* inside the comment is restricted by the topic, i.e. Peter did not give every female student in the world a book, but every female student belonging to the student group given in the topic. We consider the restriction relation between topic and comment as an anphoric binding in the sense, that the quantifier generates a domain variable to be bound by the topic, as shown in (16a). The domain variable Z is resolved in (16b) by being linked to the T-DR X.

(16) a.

K^T : X^t ; students(X)

K^C : u ; u = peter*
 y ; femal_student(y), y ∈ Z, α(Z), Z → v ; book(v), gave(u, y, v)

b.

X, u
students(X)
u=peter*

y, Z
femal_student(y)
y∈Z
Z=X
→
v
book(v)
gave(u, y, v)

5.4 Dangling Topics

There is a group of TCS frequently discussed in the literature in which the comment clause does not appear to contain any element which is anaphorically related to the topic directly or indirectly. The topics seem to be a dangling one. In this section, we take a closer look at these examples, and argue that the comment clauses in these constructions are elliptical, so that the relating elements inside the comment are missing on the surface. By ellipsis resolution we can obtain the anaphoric relation between topic and comment.

Taking the famous firefighter example (17a). It is understood as *about that fire, fortunately the fire brigade came quickly, otherwise the fire would cause an unpleasant consequence.* Well, the second part of the comment is elliptical in (17a), but every native speaker can get it. The explanation is the language usage of Chinese to avoid expressing any catastrophe clearly. The adverb *xingkui* (fortunately) is the key word of this sentence. Because the connective adverbs *xingkui...buran* (fortunately...otherwise) work in pair, in the case that the second adverb is omitted together with the clause it ushers in, the missing part is recoverable in an appropriate context or situation. But if the adverb *xingkui* (fortunately) is missing, as in (17b), the sentence has an other meaning. This difference can be shown more clearly in (18) resembling (17). While (18a) is fine, (18b) is not acceptable because the topic without the adverb *xingkui* (fortunately) becomes a real dangling topic.

(17) a. *Nei   chang   huo,   xingkui   xiaofangdui lai   de   kuai.*
       That CL   fire   fortunately fire brigade come DE   fast
       (As for that fire, fortunately the fire brigade came quickly.)

    b. *Nei   chang   huo,   xiaofangdui lai   de   kuai.*
       That CL   fire   fire brigade come DE   fast
       (That fire, the fire brigade came quickly.)

(18) a. *Nei   zuo   fangzi,   xingkui   qunian   mei   xia   xue.*
    That   CL   house   fortunately last year   not   fall   snow
    (As for that house, fortunately it did not snow last year.)

   b. **Nei   zuo   fangzi,   qunian   mei   xia   xue.*
   That   CL   house   last year   not   fall   snow

We interpret the connective adverbial clauses *fortunately A, otherwise B* as a conjunction of the fact expressed by A and a conditional relation $\neg A \rightarrow B$, i.e. if A would not take place, then an unpleasant consequence B would arise. In the case that the second part is not uttered explicitly, the adverb *fortunately* implies the would-be condition, and its consequence is reduced to a non-specified event which the speaker does not like, and which relates somehow to the topic at the beginning of the sentences. (18a) is represented as in (19) where the topic serves as the theme of the would-be event e.
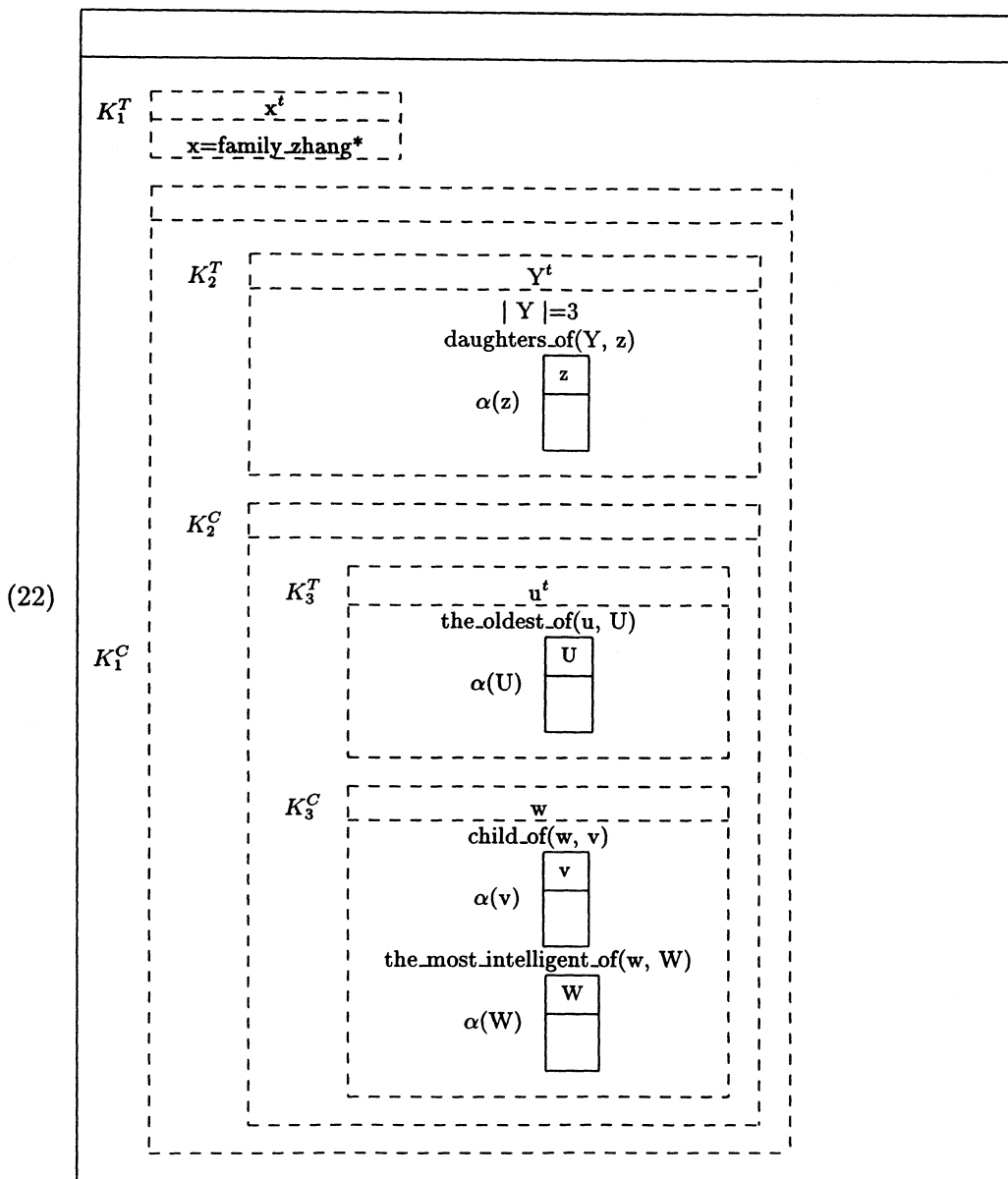
(19) a.

b.
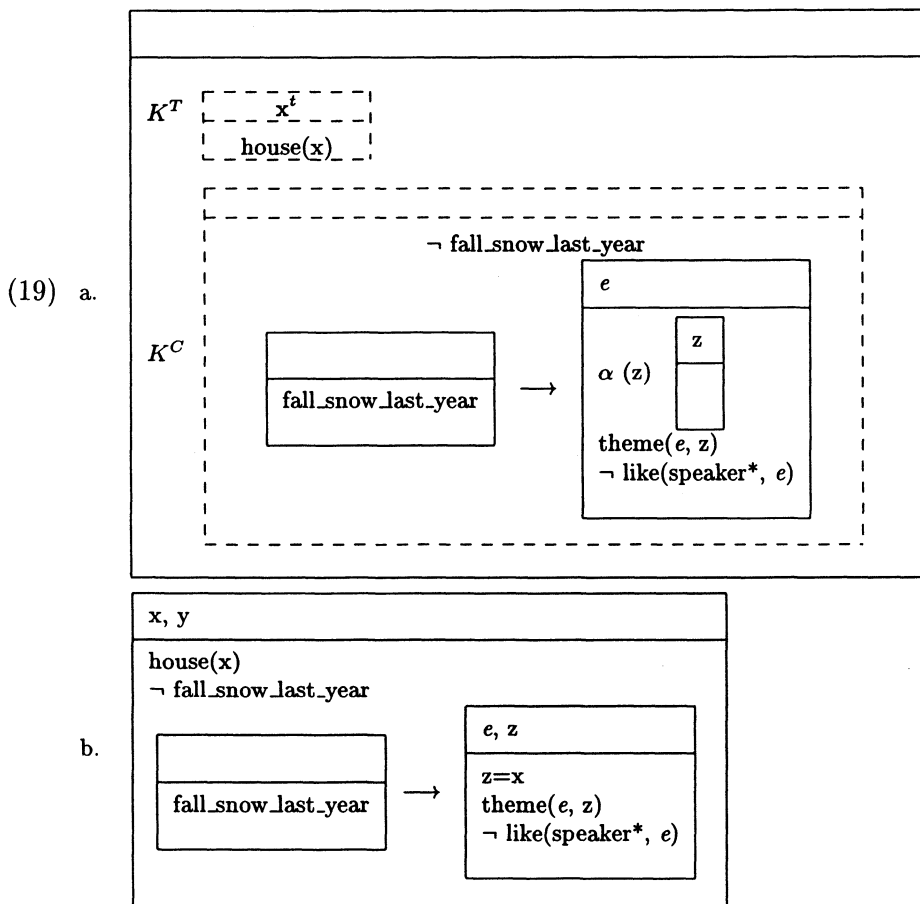
5.5 Multiple Topics

It is well known that multiple topics can occur in a TCS in Chinese. (20) for instance, contains three topics. As shown in (21), the first topic *family Zhang* subordinates the second topic *three daughters*, and this subordinates the third topic *the oldest*. The sentence means that *about family Zhang, among the three daughters of this family, the oldest of these three daughters, her child is most intelligent*.

(20) *Zhang   jia,   sange   nüer,   laoda,   haizi   zui   congming.*
    Zhang   family   three   daughters   the oldest   child   most   intelligent
    (Family Zhang, three daughters, the oldest, child is most intelligent.)

(21)

$$
\begin{array}{c}
\text{S'}\\
\text{TOPIC}_1 \qquad\qquad \text{S'}_1\\
\text{family zhang} \qquad \text{TOPIC}_2 \qquad\qquad \text{S'}_2\\
\text{three daughters} \qquad \text{TOPIC}_3 \qquad\qquad \text{S}_3\\
\text{the oldest} \qquad \text{child is most intelligent}
\end{array}
$$
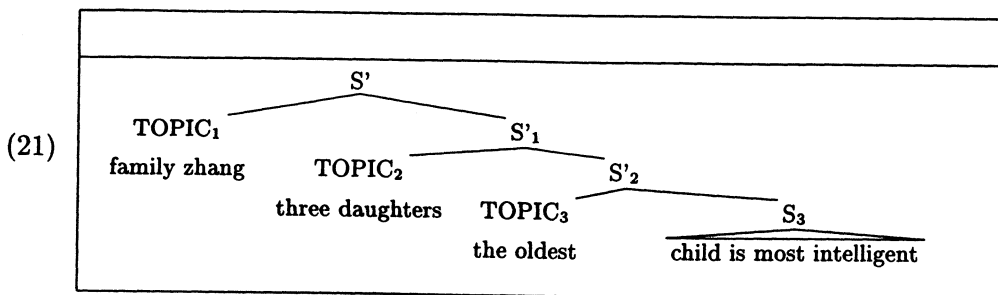
Semantically, there is a sequence of anaphoric relations between TOPIC$_1$ and TOPIC$_2$, TOPIC$_2$ and TOPIC$_3$, and TOPIC$_3$ and S$_3$. These recursive anaphoric relations can be represented in our framework as shown in (22). K_1^T, K_2^T and K_3^T are three T-DRSs. K_1^C, K_2^C and K_3^C are their corresponding C-DRSs, where K_1^C and K_2^C are TC-DRSs. The anaphoric relations are treated by resolution of TC-DRSs.

(22)

K_1^T : x^t ; $x = \text{family\_zhang}^*$

K_2^T : Y^t ; $|Y| = 3$; $\text{daughters\_of}(Y, z)$; z ; $\alpha(z)$

K_2^C :

K_3^T : u^t ; $\text{the\_oldest\_of}(u, U)$; U ; $\alpha(U)$

K_3^C : w ; $\text{child\_of}(w, v)$; v ; $\alpha(v)$; $\text{the\_most\_intelligent\_of}(w, W)$; W ; $\alpha(W)$

K_1^C

The TC-DRS resolution process proceeds from the inside to the outside. We do not illustrate the sequence of anaphor resolutions explicitly here. (23a) results by TC-DRS resolutions, and finally, the open α-DRS is resolved by the implicature that *the child of the oldest daughter of family Zhang is most intelligent of all children of the three daughters of this family*, as given in (23b).

-61-

(23) a.

x, Y, z, u, U, w, v
x=family_zhang*
| Y |=3
daughters_of(Y, z)
z=x
the_oldest_of(u, U)
U=Y
child_of(w, v)
v=u
the_most_intelligent_of(w, W)
α(W) [W |]

b.

x, Y, z, u, U, w, v, W
x=family_zhang*
| Y |=3
daughters_of(Y, z)
z=x
the_oldest_of(u, U)
U=Y
child_of(w, v)
v=u
the_most_intelligent_of(w, W)
children_of(W, Y)

6 CONCLUSION AND FUTURE RESEARCH

We have developed a semantic framework for TCS in Chinese. We argue that the semantic relation between topic and comment results from anaphora resolution. This is achieved by linking or brigding. Bridging can be seen as a repair mechanism if linking fails to meet the coherence constraint on TCS that the comment has to relate to the topic. In this way, the relation between topic and comment can be computed uniformly, as the English-style topic and the Chinese-style topic are treated as two ways of anaphor resolution.

There are some points of future research suggested by the approach presented here. The first point is that we have only considered the TCS with an NP as topic. But the syntactic category of topic is not limited to NP; VP or S can serve for topic as well ([10]). We plan to work out the semantics for TCS with non-NP topics. The second point is that the presented approach can be easily extended to topic chains ([11]). In topic chains, the topic-comment relations interact with the discourse relations between the comment clauses. We intend to investigate the discourse relations inside TCS, and to extend our analysis to capture the interaction of discourse structure with the topic-comment articulation.

References

[1] N. Asher and A. Lascarides. Bridging. Ms. University of Texas and University of Edinburgh. 1997.

[2] J. Bos and P. Buitelaar et al. Bridging as Coercive Accomodation. *Proceeding of Computational Logic for Natural Language Processing*, South Queensferry Scotland. 1995.

[3] W. L. Chafe. Giveness, Contrastiveness, Definiteness, Subjects and Topics. In Ch. Li (ed) *Subject and Topic*. pp. 27-55. 1976.

[4] H. Kamp and U. Reyle. From Discourse to Logic. Dordrecht/Boston/London. 1993.

[5] Ch. Li and S. A. Thompson. Subject and Topic: a new Typology of Language. In Ch. Li (ed) *Subject and Topic*. pp. 457-489. 1976.

[6] J. Pustejovsky. Type Corecion and Lexical Selection. In J. Pustejovsky (ed.) *Semantics and the Lexicon*. pp. 73-94. 1993.

[7] R. A. v. d. Sandt. Presupposition Projection as Anaphora Resolution. *Journal of Semantics* Vol. 9. pp. 333-377. 1992.

[8] D. Shi. The Nature of Topic Comment Constructions and Topic Chains. Ph.D Thesis at the University of Southern California. 1992.

[9] D. Westerståhl. Determiners and Context Sets. In J. v. Benthem and A. t. Meulen. (ed) *Generalized Quantifiers in Natural Language*. pp. 45-71. 1984.

[10] L. Xu and D. T. Langendoen. Topic Structures in Chinese. *Language* Vol. 61. pp. 1-27. 1985.

[11] H. Xu. Discourse Semantic Analysis of Topic-Comment Structures in Chinese. Ms. University of the Saarland. 1998.