A CENTERING THEORY ANALYSIS OF DISCREPANCIES ON SUBJECT ZERO ANAPHOR IN ENGLISH TO THAI TRANSLATION

Saranya Pathanasin¹
Wirote Aroonmanakun²

Abstract

Centering theory (CT) has been adopted in analyzing 84 zero anaphors in 50 informative texts. It is found that most zero anaphors occur in Continuation state both in English texts (source texts (ST)) and Thai translation (target texts (TT)). Zero anaphors in the TT outnumber those in ST and are found in more environments. In terms of translation, most zero anaphors in source texts remain in the same form in the target texts although some items are translated into different anaphor forms. Results indicate that zero anaphor is used to keep discourse coherence and to refer to the backward-looking center (Cb) of current utterances in both languages. Therefore, most zero anaphors in source texts are translated into zero anaphors in target texts when the CT transition state of utterances in source texts and target texts is Continuation, and are translated into other anaphors when the CT transition state in source texts is changed to another transition state in the target texts.

Constraints in translation of zero anaphors can be explained in terms of anaphor interpretation, salience of entities, syntactic constraint, and naturalness of translation. However, this paper focuses only on one type of anaphor, namely subject zero anaphor; investigation of other types of anaphor will reveal other discrepancies in using and translating anaphors from this language pair.

1. Introduction

1.1 Background of the study

This paper focuses on the use and translation of zero form as anaphor. By zero anaphor in English, we refer to the zero pronominal with subject role to a finite verb. For example:

(1) Back behind this giant reef fish’s already toothy maw looms a second set of jaws, which Ø launch from the throat, Ø grab prey from the front teeth, then Ø retreat into the dark tunnel of the eel’s esophagus.

Data shows that zero anaphor (Ø) in English can be translated into both zero and overt anaphors in Thai (target language). For example:

(2) ST³: Key to koala survival, it [bat] laps eucalyptus nectar, then Ø disperses pollen grains up to 60 miles away.

TT⁴: การมีบทบาทสำคัญในการยุติผลของการอาหารเฉพาะพืชที่นี้ทำให้นั้นคือของเดิมสีเงาและØ ช่วยกระจายของระบุ

³ ST stands for source text which in this case is English.
⁴ TT stands for target text which in this case is Thai.
Examples (2) and (3) show translation of zero anaphors in English into different anaphors in Thai, which are: zero anaphor in example (2) and a personal pronoun in example (3). Since zero anaphors in English can be translated into different anaphors in Thai, an interesting question arises: what are the factors that govern translators’ choices of anaphor when translating.

This study adopted centering theory (CT) to investigate the discrepancies between using zero anaphors in English and in Thai. It is a pioneer study that adopts centering theory in translation study. Our aims are: to analyze possible ways to translate English zero pronominals into Thai; to analyze discourse coherence in both source and target languages using centering theory; and to compare CT-transition states between the translation pair to reveal the principles that govern translators’ decisions in translating zero anaphors from English into Thai. We hypothesize that:

1) discrepancies between the uses of zero anaphor in English and Thai can be explained according to centering theory, and

2) translation of zero anaphors from English into Thai is governed by discourse structures rather than sentence structure.

To test our hypothesis, we analyzed parallel corpora of English and Thai. Samples were collected from 50 English-Thai translations, 50 source texts (ST), and 50 target texts (TT). These are informative texts selected from National Geographic

6 CT-transition state is the changes of attention in local discourse which will be described later in section 2.
magazine. This study attempts to answer the following two research questions:

1. To what extent can centering theory be used to analyze discrepancies between zero anaphor in English and Thai translation?

2. In which environments do zero anaphors remain the same and change form in translation of English to Thai?

This paper is organized as follows: Section 1, Introduction; Section 2, overviews of CT and review of particular aspects of CT adaptation in Thai; Section 3, data and research methodology; Section 4, results; and Section 5, conclusion.

2 Centering theory

2.1. CT overviews

Centering theory (CT) is formulated as a theory that relates the center of attention, choices of referring expressions, and perceived coherence of utterances, within a discourse segment (Grosz and Weinstein 1995 cited in Walker, Joshi and Prince 1998: 20). CT arose from the original work of Barbara Grosz in 1977 (Joshi and Miltsakaki 2006: 223). The centering model explains the perceived coherence of discourse by capturing the center of attention in discourse. The center of attention is a member of entities in a given discourse, and it has been seen as an interesting approach to anaphor study. Centering model can explain the different degrees of coherence in discourse as demonstrated below:

(4) 4.1) a) John went to his favourite music store to buy a piano.  
   b) He had frequented the store for many years.  
   c) He was excited that he could finally buy a piano.  
   d) He arrived just as the store was closing for the day.

4.2) a) John went to his favourite music store to buy a piano.  
   b) It was a store John had frequented for many years.  
   c) He was excited that he could finally buy a piano.  
   d) It was closing just as John arrived.  

(Joshi and Miltsakaki 2006: 224)

CT predicts that discourse (4.1) is easier to process than (4.2) because (4.1) is more coherent than (4.2). In (4.1), John is the center of attention from (a) to (d), while the center of attention shifts in discourse (4.2) between John in (a) and (c), and the store in (b) and (d).

CT provides a set of definitions, constraints, and rules to formulate the transition state in local discourse which expresses how the choice of linguistic items affects hearers’ processing (Kameyama 1998: 90). This transition state expresses the relationship between utterances in discourse which reflects the degree of coherence.

Centering theory definitions:

As Brennan, Friedman and Pollard (1987: 1) explained: A discourse segment consists of a sequence of utterances $U_1,\ldots,U_m$, with each utterance ($U_j$) associated with a list of forward-looking centers, $C(U_j)$ consisting of those discourse entities that are directly realized, or realized by linguistic expressions in the utterance. The ranking of an entity on this list corresponds roughly to the likelihood that it will be the primary focus of subsequent discourse. The first entity on this list is the preferred
center, Cp(Ui). U, actual centers, or is 'about', only one entity at a time for the backward-looking center, Cb(Ui). The backward center is a confirmation of an entity that has already been introduced into the discourse; more specifically, it must be realized in the immediately preceding utterance, Ui-1.

According to Brennan, Friedman and Pollard (1987) the set of forward-looking centers (Cf) consists of all entities that appear in the current utterance (Ui). They have different degrees of salience and therefore are ordered according to their grammatical roles as will be described later in the Cf Ranking section. The most salient member becomes the Preferred Center (Cp) which is likely to be the backward-looking center (Cb) of the next utterance.

The backward-looking center (Cb) is the entity that links the current utterance with the previous utterance (Ui-1) and is the center of attention in the current utterance (Ui). In each utterance, there is only one Cb.

The following example demonstrates how to determine utterance and its members:

(5) Ui, Brown University sociologist John Logan has pored over the melting pot in microcosm for 40 years.  
Cf: [John_Logan, Melting pot, Microcosm]  
Cb: [?] (no Cb)  
Pc: [John_Logan]

Note that an utterance in the present study is the updated unit which list of Cf member and Cp are updated. Utterance in our study consists of Subj + finite verb, so the sentence in example (5) is determined as one utterance. Then, we determine members of the Cf set in the above example by including all entities existing in the utterance Uj. Those entities are ranked according to their grammatical roles which will be described later in this section. Cb is the entity in the previous utterance (Ui-1) that is referred to in the current utterance. Because there is no previous utterance in example (5), the utterance Ui has no Cb. The most salient entity in the Cf set of the current utterance is Cp, so Cp (Ui) is the entity realized by John Logan.

The definition above describes how to determine Cf, Cb, and Cp. Next, CT provides Constraints on assigning these members as follows:

**Constraints**

For each utterance Ui in a discourse segment D consisting of utterances Uj,..., Um:

1. There is precisely one backward-looking center Cb (Ui).
2. For every element of the forward center list, Cf (Ui), must be realized in Ui.
3. The center, Cb (Ui), is the highest-ranked element of the Cf (Ui-1) that is realized in Ui.

(Brennan et al. 1987: 2)

These CT-constraints determine that: firstly, an utterance can have only one Cb; secondly, all members of Cf (Ui) are entities that exist in Ui. Note also that where the issue of realization has been argued by researchers in the field, the argument is not discussed in this paper. We apply the concept of realization as entities that exist or are directly referred to in the utterance. Thirdly, if there is more than one entity in Ui referred to in Ui, the highest ranked entity is the Cb (Ui).
Referring back to the definitions, Cf members are ranked according to their grammatical role. CT-Ranking says about this matter as follows:

**Ranking**

CT can be applied in any language. However, its universal properties are questionable. Researchers who have studied Ranking in CT agree that different languages can have different ranking which depending on the grammatical structure of the language (Walker, Iida and Cote 1994, Kameyama 1985, and Turan 1998). Cf-Ranking was originally proposed in English in which entities are ranked by their grammatical roles. The present study follows Ranking according to Grosz and Weinstein (1995: 16).

Cf ranking for English: Subject > Object(s) > others

We apply the above CT-Ranking in analyzing ST data but use a different Ranking for TT data by adopting Aroonmanakun (2000)’s Ranking for Thai which will be presented later in section 2.2.

Then, there are two CT-rules for analysis as presented here:

**Rules**

For each $U_i$ in a discourse segment $D$ consisting of utterances $U_1,\ldots, U_m$:

1. If any element of $\text{Cf}(U_{i-1})$ is realized by a pronoun in $U_i$, then the $CB(\text{U}_i)$ must be realized by a pronoun also.
2. Sequences of continuation are preferred over sequences of retaining; and sequences of retaining are to be preferred over sequences of shifting.
   
   (Grosz and Weinstein 1995: 16)

Rule 1 is generally called the ‘pronoun’ rule. Basically, if any entity of $U_{i-1}$ is referred to by a pronoun in $U_i$, $CB(U_i)$ must be referred to by a pronoun as well. For example:

(6) $U_{i-1}$ Melanee has a new skirt.
$U_i$ She likes it very much.

In the above example, a new skirt is referred to by the pronoun ‘it’, so Melanee, which is the $CB(U_i)$, is referred to by the pronoun ‘she’.

Rule 2 is about the preference of the transition state, which expresses how coherent the discourse is. CT-notion prefers Continuation to Retain, which is preferred to Smooth-shift, which is preferred to Rough-shift for the coherence of a discourse. Thus, the more Continuation, the more coherent it is.

Once, all members in utterances are determined and ranked grammatically under Constraints with respect to the rule. We can formulate the transition state between utterances.

**Centering transition states**

The transition state is a change of attentional state from one utterance to another utterance. The attentional state determines the center of attention, which may or may not be carried across utterances. Attentional states are associated with the salience of entities. Degrees of salience correspond with degrees of processing load required for anaphoric expression interpretation. In light of this, anaphors with less information, such as zero pronouns, are expected to refer to $CB$ because $CB$ is the salient entity whereas anaphors attached with more information such as noun phrases are expected to refer to less salient entities.
Transition states are classified according to the amount of change involved. This study adopts the Centering Transition proposed by Brennan et al. (1987: 3), as follows:

Table 1 Transition States

| Cb (U_i) = Cp (U_i) | CONTINUATION | SMOOTH-SHIFT |
|---------------------|--------------|--------------|
| Cb (U_i) ≠ Cp (U_i) | RETAIN       | ROUGH-SHIFT  |

Transition state preferences are: Continuation > Retain > Smooth-shift > Rough-shift

As mentioned above, transition states are classified according to the amount of change involved. Each type of transition reflects the change of Cb and Cp entities between a current utterance (U_i) and its previous utterance (U_{i-1}) as follows:

Continuation: Cb (U_{i-1}) is the same as that of U_i, and Cb & Cp (U_i) are the same entity.

Retain: Cb (U_{i-1}) is the same as that of U_i, and Cb (U_i) is not Cp (U_i).

Smooth-shift: Cb (U_{i-1}) is different from that of U_i, and Cb & Cp (U_i) are the same entity.

Rough-shift: Cb(U_{i-1}) is different from that of U_i, and Cb(U_i) is not Cp(U_i).

At this stage, we can take the utterance in example 5 above to calculate the transition state as follows:

(7) U_{i-1}: Brown University sociologist John Logan has pored over the melting pot in microcosm for 40 years.
Cf: [John_Logan, Melting_pot, Microcosm]
Cb: [?] (no Cb)
Cp: [John_Logan ]
Transition: NON U_i: Last year he sifted through U.S. census data from 1980 to 2010
Cf: [John_Logan (he), U.S._census_data]
Cb: [ John_Logan (he) ]
Cp: [John_Logan (he)]
Transition: Continuation

In example (7), the transition state is calculated from the Cb, the Cp of the current utterance and that of the previous utterance. There is no previous utterance for U_{i-1}; therefore, U_{i-1} has no transition. Then an entity realized by John Logan is not the Cb(U_{i-1}) but is the most salient element in the Cf (U_i). Therefore, John Logan is expected to be the Cb of the next utterance (U_i). John Logan is the most salient entity from the previous utterance (U_{i-1}). Therefore, it is the Cb (U_i) and it is referred to by the pronoun ‘he’ in U_i since John Logan is also the Cp(U_i), Cb(U_{i-1}) = [?], Cb ((U_i) = Cp (U_i)). The transition state at this point is then Continuation.

2.2 Centering theory in the analysis of Thai

This section turns attention to previous studies of CT in the Thai language where zero anaphor is the center of attention. The zero pronoun in Thai is an empty category, which can function in both subject and object slots in Thai sentences (Hoonchamlong 1991: 71). According to Aroonmanakun (2000) the zero pronoun is used commonly when its referent has the most focus in discourse. Normally, the antecedent of the zero pronoun is the Cb. Aroonmanakun (2000) investigated zero pronoun resolution in Thai discourse by using CT. In his study, the centering
algorithm was modified to resolve zero pronoun which antecedent is in further than immediate processing utterance \((U_{i-1})\). From the results of this study, Aroonmanakun proposed CT rule 1 for Thai as modified from Grosz and Weinstein (1995) as follows:

**Rules:**

For each \(U_i\) in a discourse segment \(U_1, \ldots, U_m\):

1. If some element of \(Cf(U_{i-1})\) is realized as a zero pronoun in \(U_i\), then so is \(Cb(U_i)\).

2. Transition states are ordered. CONTINUATION is preferred to RETAINING and RETAINING is preferred to SHIFTING. Aroonmanakun (2000: 135)

Aroonmanakun’s reforming algorithm has been proven suitable with Thai discourse segments containing a zero pronoun, and the present study adopts this algorithm in analysis.

As mentioned above, Cf ranking was found to be varied across languages. Since ranking in Thai discourse has never been proposed, this study will follow the ranking adapted by Aroonmanakun (2000) in his analysis of Thai zero pronouns. The ranking was originally proposed by Kameyama (1985) (cited in Aroonmanakun 2000) in a CT analysis of Japanese. Although Japanese and Thai are different, this ranking has been proven useful in Thai discourse analysis by Aroonmanakun as follows: [Topic > Subject > Object > Others]. The present study follows this ranking in its analysis of Thai (target texts).

### 3. Research methodology

#### 3.1 Data collection

Data in our study are parallel corpora of English (source texts) and Thai translation (target texts). Source text (ST) data was collected from 50 articles in the National Geographic magazine. All articles are informative. Each article contains 125-225 words. Target text (TT) data is the translation of ST published in National Geographic Thailand in the same issues with their ST.

The study focuses only on zero anaphor in English informative texts as rendered in Thai translation. By means of zero anaphor in the English source text, we refer to the zero pronominal with subject role to a finite verb and point to the antecedent in the previous utterance. Deictic zero pronouns are disregarded. Those in quotations are also excluded because they can be considered as unbound anaphors with links to an entity that was introduced and developed in conversation, not in the narrative of the text.

#### 3.2 Data analysis

The data was analyzed quantitatively and qualitatively. There are three steps in the analysis.

The first step is compiling the corpora. All English texts (ST) and Thai translation texts (TT) are separated into utterances. An utterance in our study consists of a Subj+finite Verb. Therefore, compound and complex sentences will be broken down into clauses. For example:\(^7\)

\(^7\) a) \(U_j^{com}\) stands for compound clause of \(U_j\)

\(^7\) b) \(U_j^{sub}\) Stands for subordinate clause of \(U_j\)

---

51
4. Results and discussion

This section presents the results of CT analysis which can reveal discrepancies between the uses of zero anaphor in English and Thai.

4.1 Comparison of zero anaphors in English-Thai corpora

Results show very similar trends of CT-transition states between English (ST) and Thai (TT). It could be said that ST and TT are equally coherent as presented in Table 2.

Then, we surveyed the distribution of zero anaphors in all 50 texts, both in ST and TT. We found in both languages that zero anaphors occur mostly in the Continuation-state, as shown in Table 3.

As can be seen from Table 3 below, zero anaphors occur mostly in Continuation-states in both languages: 75% (63 items) in ST; and 72.7% (107 items) in TT. The high number of zero anaphors in the Continuation state shows that zero anaphor is used in discourse segments in which coherence is kept. The use of zero anaphors corresponds to degrees of salience of the entity that it refers to, as will be discussed later in section 4.2.3. Following this up further, the number of zero anaphors in ST is fewer than in TT as seen in Table 3. This reveals that zero anaphors occur in TT in greater environments than in the ST. Then, the translation of zero anaphors was surveyed. The result is presented in Table 4.

The second step is the centering theory analysis. A Centering model is adopted. In all utterances in ST and TT, members of the Cf set, Cb, and Cp are determined. Then, CT-transition states between utterances are computed.

The last step is comparing CT-transition states. The transition flows in ST and TT are compared. This will reveal differences in discourse structure that govern translation of zero anaphors. The discrepancies in the translation of zero anaphors will be explained from a Centering point of view.

The second step is the centering theory analysis. A Centering model is adopted. In all utterances in ST and TT, members of the Cf set, Cb, and Cp are determined. Then, CT-transition states between utterances are computed.

The last step is comparing CT-transition states. The transition flows in ST and TT are compared. This will reveal differences in discourse structure that govern translation of zero anaphors. The discrepancies in the translation of zero anaphors will be explained from a Centering point of view.
4.2 Discrepancies of zero anaphor in ST and TT

4.2.1 General usage

Zero anaphors occur mostly in the subject position of compound sentences where the subject of the main clause and the compound clause share the same referent, and in subject position of subordinate clauses where zero anaphors in both languages refer to the nearest entity. We found that most zero anaphors in our data...
are in the Continuation-state and point to the 
Cb both in the ST and TT. For example:

(9) ST: Brown University sociologist 
John Logan has pored over the melting 
pot in microcosm for 40 years. 
Cf: [John_Logan, Melting_pot, 
Microcosm] 
Cb: [?] 
Cp: [John_Logan] 
Transition: NON

U, Last year he sifted through U.S. 
census data from 1980 to 2010 
Cf: [John_Logan (he), 
U.S._census_data] 
Cb: [John_Logan (he)] 
Cp: [John_Logan (he)] Transition: Continuation 

U_com and Ø identified 20 traditionally 
multiethnic metropolitan centers, 
including Los Angeles, Newark, and 
Houston. 
Cf: [John_Logan (Ø), 20_Multiethnic_ 
Metropolitan_Center, Los_Angeles, 
Newark, Houston.] 
Cb: [John_Logan (Ø)] 
Cp: [John_Logan (Ø)] 
Transition: Continuation 

TT: องค์ ในแก้ นักสังเคราะห์จาก 
มหาวิทยาลัยบรอัม ศึกษาปั๊มนักทอง 
วัฒนธรรม (melting pot) ในสังคมขนาดเล็กมา 
นานว่า 40 ปี 
Cf: [John_Logan, Melting_pot, 
Microcosm] 
Cb: [?] 
Cp: [John_Logan] 
Transition: NON 
[cénn0 loo0kænn0 nak3sap4khom0- 
wit3th1jaa0 caak1 ma3haa4- 
wit3th1jaa0laj0braaw0 sik1saa4 
baw2læw3m4 thaæn0 wat3tha3na3tham0 
(melting pot) nàj0 sap4khom0 
khànaal1 lek3 maa0 naan0 ruuam2 
si1sip1 pii0] 
(John Logan: sociologist from University 
of Brown; studies the melting pot in a 
small society for 40 years)

U, เมื่อปีที่ผ่านมา จัดซื้อกินข้อมูลสังคมใน 
ประชากรเสื่อม ตั้งแต่ปี 1980-2010 
Cf: [John_Logan (tha), U.S._census_ data] 
Cb: [John_Logan (tha)] 
Cp: [John_Logan (tha)] 
Transition: Continuation 

[mii2 piii0 thii2 phaan1maa0 khaw4 
siip1khon3 sam4ma3 noo0- 
pra1chaao0kænn0 sa1ha1rat3 taap2te1 
pii0 1980-2010] 
(Last year, he investigated a census of the 
U.S. data from 1980 to 2010)

U_com และ Ø ระบุผู้ถูกเลือกเพื่อนๆอาศัย 
ตั้งแต่ 20 แห่งซึ่งรวมถึงลอสแองเจลิส นิวยอร์ก 
และอีสเทิร์น 
Cf: [John_Logan (Ø), 20_Multiethnic_ 
Metropolitan_Center, 
Los_Angeles, Newark, Houston.] 
Cb: [John_Logan (Ø)] 
Cp: [John_Logan (Ø)] 
Transition: Continuation 

[le3 ø ra3bul suun4klaan0 leη1 
pha3hu1chaat2ti1phan0 daŋ2dyvm0 20 
hη1 siŋ 2 ruuam0thiη4 4ọọ3 ʔọeŋ 
0cee0lit3 niwɔsk1 le3 hiw0 sa1tan2] 
(and Ø identified centers of 20 traditionally 
multiethnic metropolitan, which included Los 
Angeles, Newark, and Houston)
4.2.2 Discrepancies in syntax

Discrepancies are found in the use of zero anaphors between English and Thai in our data. Analysis reveals that zero anaphors are used in TT in more environments than in ST due to the fact that Thai is a pro-drop language which allows subject and object omission (Hoonchamlong 1991). In Thai the subject of the main clause can be omitted. It is found only in TT that the zero anaphors occur in the initial position of sentences when sentences are in the Continuation-state; this aspect is not found in ST data. For example:

(10) ST: If it success....  
    TT: หาก Ø สั่งขึ้นเจรจ...  
    haak1 ø sam4ret1...  
    (If Ø success....)

ST : Or do they?  
TT: แต่ Ø เป็นเช่นนั้นจริงหรือ  
    [te1 ø pen0 chen2nan3  
    cin0rii4]  
    (but Ø be that true Y/N)

Interestingly, zero anaphors in the initial position of sentences in the TT can occur only in the Continuation-state in our data. It is not found in other CT-transition states. We assume that this finding is a phenomenon of informative texts.

4.2.3 Zero anaphors and salient entities

It has been found that zero anaphors commonly point to the Cb which is the most salient entity in the utterance. With regard to this matter, Givon (1983: 359) demonstrated a scale of correlation between the degree of continuity/predictability of topic NP's and the average size of the marking devices used to express them, which start from zero pronoun for the most continuous/predictable topic to modified DEF-NP when the degree is decreased. Although Givon’s scale of Topic Continuity/Predictability and Marking Size was the result of a spoken English analysis, it corresponds well with the result of our CT-analysis both in ST and TT where zero anaphors are used to refer to the most salient entity which is continually predicted to be the topic.

As the result recorded in Table 4 above shows, zero anaphors in the ST are translated into zero anaphors in the TT when utterances of the ST and TT are in the Continuation state and when they refer to the Cb as shown in example (9) above. On the other hand, when there are changes of salient entity and transition state in the ST and TT, we found that zero anaphors in the ST are translated into other anaphors in the TT. For example:

(11) ST: (U) The breeders will cross those cattle to retain the pertinent DNA, jettison the rest, and make bovines that,

    Cf: [breeder, cattle, DNA, bovines]  
    Cb: [breeders]  
    Cp: [breeders]  
    Transition: Continuation

    (U[sub]) in about a decade, Ø are expected to look and act just like their extinct ancestors.

    Cf: [bovines (Ø), Aurochs (their extinct ancestors)]  
    Cb: [bovines (Ø)]  
    Cp: [bovines(Ø)]Transition: Smooth-shift

    TT: U, จากนั้นพวกเขาจะทำการผสมข้ามพันธุ์ระหว่างวัวเหล่านี้อีกครั้งเนื่องจากเวลา 10 ปี
4.2.4 Double zero anaphors

Findings confirm the modification of CT rules for Thai discourse as proposed by Aroonmanakun (2000) which state that: If some element of Cf(Ui) is realized as a zero pronoun in Ui, then so is Cb(Ui). It was found in the TT that if an entity of Cf(Ui) is referred to by a zero anaphor in the current utterance, Cb(Ui) is referred to by a zero pronoun as well. For example:

(12) ST: Ui. Popular in Alpine villages centuries ago, Krampus scared kids straight-his long red tongue upped the fear factor-
Cf: [demon (Krampus), kids, long_red_tongue, Alpine_villages]
Cb: [demon]
Cp: [demon]
Transition: Continuation

TT: Ui, and Ø taught them that evil bows before good.
Cf: [demon (Ø), kids(them)]
Cb: [demon (Ø)]
Cp: [demon(Ø)]
Transition: Continuation

In example (11), the Cb (Ui) was changed from bovines in the ST to breeders in the TT caused by the change of passive voice to active voice. The translator converted voices to be suitable to Thai discourse, resulting in not only the change of salient entity, but also transition state in the Ui. This result suggests that the zero anaphor in the above example is translated into an NP because of the change in the Cb.
4.2.5. Zero anaphor on discourse level

Following this further, we found in the TT that when an entity is continued as the most salient in a discourse segment, the zero anaphor is expected to be its anaphoric reference. On the other hand, when the status of the salient entity is not carried over in the discourse segment, other anaphors are expected. For example:

(14) ST: \( U_i \) Together they report
Cf: [Researcher_groups (they)]
Cb: [Researcher_groups (they)]
Cp: [Researcher_groups (they)]
Transition: Smooth-shift

\( U_i \) that the female of this solitary bee-which eschews hive life-digs a shallow tunnel in loose ground with room for one or two chambers, or brood cells, each up to two inches deep.

Cf: [Bee, Tunnel, Chamber, Cells]
Cb: [?]
Cp: [Bee]
Transition: NON

\( U_j \) She then papers the cell walls with overlapping petals flown in one by one from nearby fields,

Cf: [Bee (she), Cell_Walls, Petals, Fields]
Cb: [Bee (she)]
Cp: [Bee (she)]
Transition: Continuation

\( U_j \) \( \emptyset \) gluing two layers together with a thin coat of mud.

Cf: [Bee (\( \emptyset \)), Layers, Mud]
Cb: [Bee (\( \emptyset \))]
Cp: [Bee (\( \emptyset \))]
Transition: Continuation
(They reported similarly that)

These female bees do not like making hives like other bees)

(Each room deep about one to two inches)

After that it will bring flower petals to glue together on wall rooms)

(by Ø fly to collect from field nearby one by one)

Cf : [Bee (Ø), Fields, Petals]
Cb : [Bee (Ø)]
Cp : [Bee (Ø)]
Transition: Continuation

[dooj0 Ø bin0 paj0 kep1 maa0 caak1 thuen2ja2 klaj2 khian0 thii0la3kliip1]

U_k klub โดย Ø นินไปเก็บมาจากทุ่งหญ้าใกล้ถึงที่ละกี้บ

Cf : [Bee (Ø), Mud, Petals]
Cb : [Bee (Ø)]
Cp : [Bee (Ø)]
Transition: Continuation
We found that zero anaphors can occur in the initial position of Thai sentences in the Continuation state because Thai is a pro-drop language. On the other hand, zero anaphors in English cannot occur in this environment. In addition, our results contribute to the previous study by Aroonmanakun (2000) regarding zero anaphors in Thai. It is only in the TT that an utterance can contain two zero anaphors whereas an utterance in the ST can have only one zero anaphor. In terms of translation, we found that zero anaphors in the ST are translated into zero anaphors in the TT when an utterance is in the Continuation state and into overt anaphors when the transition state in the ST is changed in the TT.

Not only differences in the use of zero anaphors between English and Thai, but also differences in discourse between the two languages affect discrepancies in the translation of zero anaphors. Here, we have analyzed and identified discourse factors that govern translators in using and translating zero anaphors. These factors, i.e., meaning and anaphor interpretation; syntactic constraints; and the naturalness of the target language, are included in the following discussion.

4.3.1 Meaning and interpretation

We found that zero anaphors in the ST tend to be translated into zero anaphors in the TT by default as shown in Table 4. This direct translation method is employed to keep discourse coherence. However, it was found that some items are translated into overt anaphor forms which are: personal pronoun and definite NP as shown in Table 4. At this point, zero anaphors in English can be translated into different anaphor forms in Thai. Analysis shows that when working on zero
anaphors, translators take into account anaphoric interpretation and ambiguity. Which form of anaphor is chosen relies on the basic principle that readers must be able to interpret its antecedent conveniently and unambiguously without increasing the processing load.

We found that change in the CT transition state and salience entity is directly relevant to the change in zero anaphors between the ST and TT. Thus, overt anaphors are employed in the TT to help Thai readers interpret referents conveniently when the center of attention is shifted between $U_{i-1}$ and $U_i$. Some good examples on discrepancies in the translation of zero anaphors are presented in Examples 11 and 14 above.

### 4.3.2 Syntactic constraint

The uses of zero anaphors in English and Thai follow the syntactic constraints of each language. Automatically, the translation of anaphors from English to Thai is governed by syntactic constraints in the Thai language. An important aspect to discuss here is that Thai is a pro-drop language which allows subjects to be omitted while English is not. Consequently, our data showed that the zero form in the TT occurs in higher numbers than in the ST, and is found in more environments of use. An example of these discrepancies is presented in Example 10 above. It can be assumed that the discrepancy in the translation of zero anaphors is under syntactic constraints.

Another syntactic constraint that governs the uses and translation of zero anaphors in our data concerns to possessive pronouns. While authors of ST, have to link inalienable possession with possessor by a possessive pronoun, such anaphors cannot be omitted. For example:

(15a) ST: To go faster, they (snakes) shift their weight by slightly raising parts of their body, as we do.

In the above example, the possessive pronoun ‘their’ in the ST links the possessor snakes with the body part which is considered an inalienable possession. In English, the possessive pronoun is needed as a linkage between weigh and body with snakes. On the other hand, a possessive pronoun is not necessary in Thai when the possession is inalienable (i.e. body part). Therefore, the translator translates the possessive pronoun in this environment into a zero anaphor. Moreover, if possessive pronouns were kept, it would create redundancy. The above sentence is translated as:

(15b) TT: พวกเขาเคลื่อนไหวเร็วขึ้น พวกมัน (snake) จะย้ายน้ำหนักโดยขยับส่วนก้น

[haa1k1 tɔŋ2kaan0 liiaj3 haj2 rew0khin2 phuak2man0 ca1 thaaj1 naam3nak1 ø
do0j0 jok3tua0 khin2 ø lek3nɔɔj3]

(To go faster, they shift weight by slightly raising parts of body)

### 4.3.3 Naturalness of language in the target texts

In attempting to keep discourse coherence, translators must be aware of the naturalness of language. For example, we found that translators convert passive voice in the ST into active voice in the TT as it is generally accepted that passive voice can sound unnatural in Thai as seen in Example 11 in section 4.2. Moreover, we found that translators often combine/rearrange a discourse segment of
several utterances in the ST into a coordinate sentence in the TT and employ a zero anaphor in the subject position of a subordinate clause to make the TT sound natural in Thai. For example:

(16) ST: Today cranberries are marketed year-round in both juice and dried form. They’re also touted as a health food, because they can keep bacteria from clinging to the urinary tract and Ø may even play a role in cancer prevention.

TT: ปัจจุบันสามารถตลาดนี้พืชที่มีบุกเบิกและแบ่งปันและ Ø ต้องยืนยันอาหารสุขภาพ เนื่องจาก Ø มีสารพัดลักษณะที่จัดเก็บความคงตัวและความและ Ø อาจเข้าไปยังกลุ่มโรคป่าต้นน้ำคือโดย

(pat1cu1ban0 raw0 rap3pra1thaan0 khreen0byy0ori2 daaj2 tal1co1 than3pii0 than3 naj0 beep1chiiai2 le3 beep1heep2 le3 o jan0thii4pen0 ?aa0thaan4 suk1kha1phaap2 nihan2caak1 o mii0 sap1pha0khun0 p3en2kan0 maj2haj2 bek1thii0ria0 k3l than1dyy0nu0pat1sa0wa3 le3 o ?aat1chuaj2 p3en2kan0 rook2 ma3renj0 ?iik1duaj2] (at present, we can eat cranberry all year both in juice and dried and Ø is accepted as health food because Ø can prevent bacteria from clinging to the urinary tract and Ø may help prevent cancer as well.)

In the above example, the ST sentence is rearranged in the TT. The translator combines two sentences in the ST into one coordinate sentence in the TT. Then she employs three zero anaphors in subject position to keep the entity realized by the dart berries (cranberries) in focus. This change is to make the TT sound natural in Thai.

5. Conclusion

In adopting centering theory in the analysis of discrepancies in the uses and translation of zero anaphors between English and Thai, the study found that zero anaphors point to Cb(U), when U is in the Continuation-state in both languages. Zero anaphors are used in the TT in more environments than in the ST due to syntactic constraints—especially the status of Thai as a pro-drop language. In addition, zero anaphors in English can be translated into different anaphor forms in Thai. Our findings reflect that different language structures influence translators to rearrange discourse segments for the sake of naturalness in the target language. Consequently, the rearrangement affects the salience of entities in utterances and CT-transition states. These cause discrepancies in the translation of zero anaphors from English into Thai.

The results can answer our two research questions stated in the introduction. That is to say: we can adapt centering theory in the analysis of discrepancies in uses of the zero anaphors between English and Thai; and the discourse structure of the TT governs the rearrangement of the target language, and consequently, affects the translation of zero anaphors.

Hopefully, the results of the study can benefit researchers in translation and discourse study. For further studies, CT analysis of other anaphor types such as
demonstrative pronouns, personal pronouns, and definite NPs should be conducted. Studies on discrepancies of anaphors in other translation genres, and in larger sample sizes, should confirm the results of the present study, and reveal other discrepancies in English to Thai translation. It is also interesting to see if zero anaphors can occur in the initial position of a sentence in other CT-transition states besides Continuation, as reported in our study. Lastly, research on Cf-ranking in Thai should be conducted for further application of centering theory in Thai.

Acknowledgements

This project is financially supported by the 90th Anniversary of Chulalongkorn University Fund (Ratchadaphiseksomphot Endowment Fund) and partially funded by The National Research University Project of CHE (HS1153A). Our grateful thanks are also extended to Assistant Professor Dr. Rachada Pongprairat for her assistance in Thai transcription. Special thanks to anonymous reviewers for helpful discussion and comments throughout our work.

References

Aroonmanakun, Wirote. 2000. Zero Pronoun Resolution in Thai: A Centering Approach. Paper presented at The International Conference on Human and Machine Processing of Language and Speech. National Electronic and Computer Technology Center. Bangkok.

Brennan, Susan, Marilyn W. Friedman and Carl J. Pollard. 1987. A Centering Approach to Pronouns. Paper presented at The 25th Annual Meeting of the Association of Computational Linguistics, 6-9 July, 1987, Stanford.

Givon, Talmy. 1983. Topic Continuity in discourse: A Quantitative Cross Language Study, Amsterdam/Philadelphia: J. Benjamins.

Grosz Barbara J. 1977. The representation and use of focus in dialogue understanding (technical report no. 151). Menlo Park CA: SRI International.

Grosz, Joshi and Scott Weinstein. 1995. Centering: A Framework for Modelling the Local Coherence of Discourse. Computational Linguistics 21:203-255.

Hoochchamlong, Yuphaphan. 1991. Some Issues in Thai Anaphora: A Government and Binding Approach. Doctoral dissertation, University of Wisconsin, Madison.

Joshi, Aravind and Eleni Miltsakaki. 2006. Anaphora Resolution: Centering Theory Approach. In Encyclopedia of Language & Linguistics (Second Edition), edited by Keith Brown, pp. 223-230. Elsevier Science.

Kameyama, Megumi. 1985. Zero Anaphora: The Case of Japanese. Doctoral dissertation, Stanford University.

---. 1998. Ranking Forward Looking Centers. In Centering Theory in Discourse, edited by Waler et al., pp. 90-112. Oxford: Clarendon Press.
Turan, Umit D. 1998. Ranking Forward Looking Centers in Turkish: Universal and Language-Specific Properties. In Centering Theory in Discourse, edited by Waler et al., pp. 139-160. Oxford: Clarendon Press.

Walker, Marilyn, Sharon Cote and Masayo Iida. 1994. Japanese discourse and the process of centering. Computational Linguistics 20.2: 193-232.

Walker, Marilyn, Aravind Joshi and Ellen Prince. 1998. Centering Theory in Discourse. New York: Oxford University Press.