

ANAPHOR RECONSTRUCTION IN THAI RELATIVE CLAUSES

Yunchuan Chen
Duke University
yunchuan.chen@duke.edu

Abstract

Whether an anaphor inside the head of relative clauses can take the embedded subject as its antecedent is commonly used to test the head derivation in relative clauses (e.g., Bhatt 2002; Schachter 1973). This paper uses a truth value judgment experiment to investigate whether the anaphor *tuaeng* ‘self’ within the head noun phrase of Thai relative clauses can be bound by the embedded subject. The results suggest that such binding is prohibited, which further indicates that the head noun phrase of Thai relative clauses must be base-generated external to the modifying clause.

Keywords: Thai, relative clauses, anaphor, reconstruction
ISO 639-3 codes: tha

1 Introduction

There have been three main analyses for relative clauses (RCs): the head raising analysis, the head external analysis and the matching analysis. The three analyses are different in two aspects: (i) how the head of RCs originates and (ii) how the gap is identified with the head of the RC.

First, the head raising analysis is also known as the promotion analysis (e.g., Brame 1968, Schachter 1973), which states that the head noun phrase (NP) is generated inside the RC and then raised out of the RC. The example in (1) illustrates the head raising process proposed by Kayne (1994).

(1) [DP the [CP cake_i [C that [IP John ate t_i]]]]

In contrast to the head raising analysis, the head external analysis claims that the head NP is base-generated external to the RC, where there is a relative operator moving to [Spec, CP] (e.g., Chomsky 1977; Jackendoff 1977).

(2) [DP the cake_i [CP op_i/which_i John ate t_i]]

For the matching analysis, according to Sauerland (1998) and Hulsey and Sauerland (2006), the head NP is base-generated external to the RC, but there is a phonologically covert internal head inside the RC, equivalent to the external head. The internal head is deleted under identity with the external head, as in (3).

(3) [DP the [cake]_i; [CP [op/which ~~eake~~]_i; John ate t_i]]

To summarize, the three analyses are different as to how the head NP is derived. While the head raising analysis claims that it is raised from within the RC, under the head external analysis and the matching analysis, it is base-generated external to the RC. Although we want an analysis that can account for all languages, RCs seem to be subject to one or more of the above analyses across languages and even within the same language (Aoun & Li 2003).

2 Reconstruction effects of the head noun phrase in relative clauses

Whether the head NP of RCs can reconstruct into and get interpreted inside the RC at Logical Form (LF) has long been considered to be a diagnostic to determine the head derivation of RCs in a language (See Aoun & Li 2003 for Lebanese Arabic and Chinese; Miyamoto 2017; Murasugi 2000 for Japanese; Han 2013 for Korean). Under Chomsky's (1993) proposal that reconstruction occurs only when syntactic movement is involved, the head NP should be raised from within the RC if reconstruction effects are observed. In contrast, if there are no such effects, the head NP should be base-generated external to the RC. Moreover, due to the possibility of having an internal copy of the head NP inside the RC, reconstruction is also predicted to be possible under the matching analysis. Thus, evidence concerning reconstruction does not differentiate between the head raising analysis and the matching analysis (Bhatt 2002). Nevertheless, if no reconstruction effects are found, the head external analysis is surely supported.

3 The head raising analysis for Thai relative clauses

With respect to the reconstruction effects of the head NP in Thai RCs, Jenks (2011) provided idioms and adjectival modifiers as two pieces of evidence to argue that the head NP is raised from within the RC.

3.1 Idioms

One diagnostic that is often used to examine the head derivation in RCs is based on idioms. Schachter (1973) provided such evidence to argue that the head NP of English RCs should be raised from within the RC.

- (4) She is keeping careful track of her expenses.
 (5) *The careful track pleases me.
 (6) The careful track that she's keeping of her expenses pleases me.

The expression *keep track of* in (4) is an idiom in English, meaning 'to monitor.' Part of the idiom cannot be used independently, as evidenced by the ungrammatical status of (5): *careful track* cannot stand alone and has to be interpreted together with other parts of the idiom. However, when the *careful track* in (4) is relativized, as in (6), it can maintain its idiomatic interpretation. Under the assumption that all elements of an idiom are generated locally as a whole unit, Schachter (1973) pointed out that *track* in (6) has to reconstruct into the RC in order to receive its idiomatic meaning. Thus, the head of the RC in (6) must have been raised from within the RC.

With the same diagnostic, Jenks (2011) used Thai idioms to argue that the head NP of Thai RCs should be raised.

- (7) *lukmai lon mai klai ton*
 nut fall not far tree
 'Lit. The nut doesn't fall far from tree.'
 'Children aren't that different from their parents' (Jenks 2011:150)

Jenks stated that when the subject of the idiom, *lukmai* 'nut,' is relativized, the idiomatic interpretation still remains, as in (8).¹

- (8) *lukmai thi lon mai klai ton thamhai phomae sabaichai*
 nut that fall not far tree let parents content
 'Children that aren't different from their parents put their parents at ease.' (Jenks 2011:150)

Since the head NP *lukmai* 'nut' can be interpreted with the other parts of the idiom inside the RC, the head NP should be able to reconstruct into the RC, which in turn supports the head-raising analysis² for Thai RCs. However, there are many other examples where the idiomatic interpretation cannot be maintained when part of the idiom gets relativized. Now we look at two examples and compare them with their counterparts in Chinese, whose RCs have been argued to involve head raising (Aoun & Li 2003; Huang, Li & Li 2009). The

¹ Based on my consultation with several native Thai speakers, their judgments vary on the grammaticality of (8).

² As discussed, the matching analysis should also be possible.

first example is from the expression *kin bunkao* ‘to eat past good deeds,’ which idiomatically means ‘to live off old wealth or inheritance.’ When we relativize the NP *bunkao* ‘past good deeds,’ the idiomatic meaning is not available, as in (9).

- (9) *khao mai mi ngan tae [bunkao [thi khao kin dai]] mi makmai.*
 he not have job but past good deeds that he eat can have a lot
 ‘Lit. He does not have a job but there are a lot of past good deeds that he can eat.’
 *‘He does not have a job but there is a lot of old wealth that he can live off.’

The awkwardness of (9) cannot be attributed to the other parts of the sentence because, if we replace *bunkao* ‘past good deeds’ with *ahan* ‘food,’ the sentence becomes totally natural, as in (10).

- (10) *khao mai mi ngan tae ahan thi khao kin dai mi makmai.*
 he not have job but food that he eat can have a lot
 ‘He does not have a job but there is a lot of food that he can eat.’

Interestingly, there is a similar idiom in Chinese, *chi laoben* ‘eat old money,’ idiomatically meaning ‘to live off one’s old wealth or skills.’ However, in Chinese, when the NP *laoben* ‘old money’ is relativized, the idiomatic interpretation can be maintained. The Chinese equivalent of (10) is (11).

- (11) *ta mei you gongzuo, dan [ta keyi chi de] laoben] henduo.*
 he not have job but he can eat DE old money a lot
 ‘Lit. He does not have a job but there is a lot of old money that he can eat.’
 ‘He does not have a job but there is a lot of old wealth that he can live off.’

In (11), the head NP *laoben* ‘old money’ can be interpreted with the embedded predicate *chi* ‘eat’ to generate the idiomatic meaning, which suggests that it can reconstruct into the RC at LF.

Another example is from the expression *kin namta* ‘to drink tears’ in Thai, which idiomatically means ‘to be in sorrow/to cry.’ The example in (12) shows that when the NP *namta* ‘tear’ is relativized, the idiomatic interpretation is not available.

- (12) *thuk khon ru namta thi khao kin*
 every person know tear that he eat
 ‘Lit. everyone knows the tear that he ate.’
 *‘Everyone knows the sorrows that he experienced.’

Since *namta* ‘tear’ cannot be interpreted with the embedded predicate *kin* ‘to eat,’ it should not reconstruct at LF. Moreover, this idiom *kin namta* ‘to drink tears’ also has a similar counterpart in Chinese: *yanxia leishui* ‘swallow tears,’ which idiomatically means ‘to endure sorrows and pretend to be happy in front of others.’ Contrary to Thai, when the NP *leishui* ‘tear’ is relativized to be the head of an RC, the idiomatic meaning is still available, as in (13).

- (13) *mei-gen-ren dou zhidao [[ta yanxia de] leishui].*
 every-CL-person all know he swallow DE tear
 Lit. ‘everyone knows the tears that he swallowed.’
 ‘Everyone knows the sorrows that he endured.’

In (13), the head NP *leishui* ‘tear,’ along with the embedded predicate *yanxia* ‘swallow,’ can generate the idiomatic meaning, which implicates head reconstruction. The Chinese examples (11) and (13) are compatible with Aoun and Li’s (2003) head raising analysis for Chinese RCs. If Thai RCs also involve the same head raising mechanism, it would be hard to explain why the idiomatic meaning is unavailable in (9) and (12). Thus, the above contrasts between Thai and Chinese suggest a fundamental difference between the two languages in terms of the head derivation in RCs: if the head NP is raised in Chinese, it cannot be subject to the same strategy in Thai.

3.2 Adjectival modifiers

The second evidence that Jenks (2011) used to support the head raising analysis for Thai RCs is from adjectival modifiers, which Bhatt (2002) proposed as another diagnostic to test the head derivation in RCs. Bhatt observed that adjectival modifiers such as *first* may be interpreted in various positions inside the RC.

- (14) the first book that John said Tolstoy had written
 ‘High reading’: the first book that John ever said Tolstoy wrote (the order of saying matters)
 ‘Low reading’: the first book that John said Tolstoy ever wrote (the order of writing matters)
 (Bhatt 2002:58)

The sentence in (14) has two possible interpretations, which Bhatt called ‘high reading’ and ‘low reading.’ For the high reading, the adjectival modifier *first* is interpreted with the matrix predicate *said* while for the low reading, it is interpreted with the embedded predicate *wrote*. The availability of the low reading was argued to implicate a copy of the head NP inside the RC, which means that the head NP originates within the RC and can reconstruct into the RC at LF. Following this logic, Jenks (2011) argued that the head NP of Thai RCs should also be raised from within the RC because, as in his Thai example (15), the low reading is possible.

- (15) *bothwam chabap sutthai thi Nit phut wa Chomsky khian chue wa on phases.*
 paper CL last that Nit say that Chomsky write call that on phases
 ‘The last paper that Nit said that Chomsky wrote is “On Phases.” (Jenks 2011:151)

In (15), *sutthai* ‘last’ can be interpreted with either the matrix predicate (high reading) or the embedded predicate (low reading).

However, Heycock (2005) argued against Bhatt’s proposal by pointing out that the low reading of the adjectival modifier may not be derived from reconstruction. First, she stated that the low reading of ordinals such as *first* are only available with a limited number of embedded predicates. For instance, while the predicate *say* in (14) allows the low reading, predicates such as *know* do not, as in (16).

- (16) This is the first book that John knew Tolstoy had written
 ‘High reading’: this is the first book that John ever knew Tolstoy had written
 *‘Low reading’: this is the book that John knew Tolstoy had written first before others.

In addition, Heycock argued that if the low reading is derived from reconstruction, we predict it to co-occur with other reconstruction effects. However, this is not always true, as in (17).

- (17) This is the first picture of himself_i that Mary mistakenly thought John_i would publish.
 ‘High reading’: this is the first picture of himself that Mary mistakenly ever thought John would publish
 *‘Low reading’: this is the picture of himself that Mary mistakenly thought John would publish first.

In (17), the anaphor *himself* can be bound by *John*, which indicates that the head NP, along with its adjectival modifier *first*, has to reconstruct into a position lower than *John* at LF. But the low reading ‘this is the picture of himself that Mary mistakenly thought that John would publish first’ is not available. Why can *himself* reconstruct but *first* cannot? This is a question needed to be addressed if we adopt the view that the low reading is derived from head reconstruction. Thus, it is unclear whether the low reading is indeed generated via reconstruction. Heycock (2005) claimed that the low reading only occurs with the RC that involves a negative entailment. Bhatt and Iatridou (2012) also noted that the low reading might not be derived from head reconstruction, but from some other semantic-pragmatic mechanism. Therefore, adjectival modifiers may not be a valid diagnostic to use when we examine the head derivation of RCs.

Given the above issues with idioms and adjectival modifiers being used as diagnostics to test the head derivation of RCs, we need to look for other alternatives.

3.3 Anaphor binding

Another commonly used diagnostic to examine the head reconstruction of RCs is anaphor binding (e.g., Bhatt 2002; Schachter 1973). Specifically, we want to see whether an anaphor inside the head of RCs can take the RC subject as its antecedent. The example in (18) demonstrates that the anaphor *himself* within the head NP can be bound by the RC subject (Schachter 1973).

(18) [[The portrait of himself]_i that John_i painted t_j] is extremely flattering.

Under the assumption that English anaphors are subject to Condition A (Chomsky 1981, 1986), since the anaphor *himself* in (18) is not c-commanded by the RC subject on the surface, the head NP must reconstruct into the RC at LF so the anaphor can be c-commanded by *John*. This reconstruction further suggests that the head NP is raised from within the RC.

As for Thai RCs, Jenks (2011) stated that the anaphor binding data he collected from native Thai speakers showed substantial variation. Also, he argued that since the Thai anaphor *tuaeng* ‘self’ has logophoric uses, it would be difficult to determine whether it is a logophor or an anaphor if it can be co-indexed with the RC subject.

(19) [rupphap khong tuaeng]_{ji} [thi Daisy_j at e_i]
 photo GEN self that Daisy compress
 ‘the self’s photo that Daisy compressed’

For instance, in (19), if the anaphor *tuaeng* within the head NP can be co-indexed with the RC subject *Daisy*, we cannot conclude that the head NP can reconstruct into the RC because *tuaeng* can be either an anaphor or a logophor. If it is a logophor, it can be co-indexed with *Daisy* without being bound by it. Moreover, Jenks quoted Lasnik’s (1989) statement that Thai lacks Condition C of the binding theory.³ Thus, it should be possible for *Daisy* to be bound by *tuaeng* even if the head NP is base-generated external to the RC.

However, if *tuaeng* is found to be unable to refer to the RC subject, we can then conclude that the head NP should not reconstruct into the RC. Since Jenks’ (2011) consultation with native Thai speakers showed a substantial variation among their judgments on anaphor binding, this paper used a controlled truth value judgment experiment (Crain & Thornton, 1998) to address this issue.

4 Research Questions

A truth value judgment experiment was used to investigate the following research questions.

(20) Research Question 1: Can an anaphor *tuaeng* ‘self’ inside the head NP of Thai RCs take the subject of the RC as its antecedent?

Research Question 2: If it can/cannot, what are the implications for the head derivation in Thai RCs?

5 Experiment

Chen (2019) created a picture-matching truth value judgment experiment to investigate a similar issue in Chinese and Japanese: whether the anaphor within the head NP of RCs can refer to the RC subject in the two languages. The experimental results suggested a clear contrast between Chinese and Japanese: the anaphor can refer to the RC subject in Chinese but not in Japanese. The findings are compatible with the existing analyses that the head NP of RCs is raised from within the RC in Chinese (Aoun & Li 2003; Huang et al. 2009) but is base-generated external to the RC in Japanese (e.g., Fukui & Takano 2000, Murasugi 2000). The current Thai experiment was created based on Chen’s and every experimental stimulus was translated to Thai

³ This statement should be interpreted carefully. As pointed out by one reviewer, Condition C is violable only when the two nominals in question are exact copies of each other (Hoonchamlong 1991) and the two identical nominals are unmodified by determiner-like elements such as classifiers or demonstratives (Deen & Timyam 2018). So it is not the case that Thai is not subject to Condition C.

as closely as possible,⁴ in order for the results to be later compared with Chen’s Chinese and Japanese data. The next issue is whether the Thai anaphor *tuaeng* resembles the Chinese anaphor *ziji* and the Japanese anaphor *jibun*.

5.1 *Tuaeng, Ziji & Jibun*

Tuaeng, Ziji and Jibun are the most representative anaphors in Thai, Chinese and Japanese and they are all morphologically simplex.⁵ Interestingly, they share many properties that do not exist in other languages such as English.

First, the three anaphors lack specification of phi features including person and gender.⁶ As in (21a)-(21c), they can take the first person pronoun *wo/watashi/chan* ‘I’ or proper nouns like *John* or *Mary* as their antecedents, which means that they do not have to agree with the antecedent in person and gender features.

(21a) *wo_i/John_i/Mary_k piping-le ziji_{i/j/k}.* (Chinese)
I/John/Mary blame-PST self
‘I_i/John_i/Mary_k blamed myself_i/himself_j/herself_k.’

(21b) *watashi_i/John_j/Mary_k-ga jibun_{i/j/k}-o seme-ta.* (Japanese)
I/John/Mary-NOM self-ACC blame-PST
‘I_i/John_i/Mary_k blamed myself_i/himself_j/herself_k.’ (Aikawa 2002:157)

(21c) *chan_i/John_j/Mary_k thot tuaeng_{i/j/k}.* (Thai)
I/John/Mary blame self
‘I_i/John_j/Mary_k blamed myself_i/himself_j/herself_k.’

Second, the three anaphors allow long-distance binding. As shown below, they can be bound by the subject NP *John*, which is located outside of the clause in which the anaphor occurs.⁷

(22a) *John_i shuo Bill_j biaoYang-le ziji_{i/j}.* (Chinese)
John say Bill praise-PST self
‘John_i said that Bill_j praised self_{i/j}.’

(22b) *John_i-ga Bill_j-ga jibun_{i/j}-o home-ta to it-ta.* (Japanese)
John-NOM Bill-NOM self-ACC praise-PST that say-PST
‘John_i said that Bill_j praised self_{i/j}.’

(22c) *John_i klao wa Bill_j yokyong tuaeng_{i/j}.* (Thai)
John say that Bill praise self
‘John_i said that Bill_j praised self_{i/j}.’

⁴ There are some Chinese/Japanese verbs that do not have equivalents in Thai so they were replaced with other common Thai verbs.

⁵ They are morphologically simplex compared to their morphologically complex counterparts such as *tuakhaoeng* ‘himself,’ *tuathoeeng* ‘herself,’ *taziji* ‘himself/herself,’ *jibun-jishin* ‘self-self.’ However, *tuaeng* is not morphologically minimal. As Jenks (2011) points out, this anaphor should be composed of two independent elements: *tua*, synonymous with the classifier for animal, and *eng*, meaning something self-referential. This is supported by morpheme order of the complex anaphors *tuakhaoeng* ‘himself’ and *tuathoeeng* ‘herself,’ where *tua* can be split off from *eng*. Further, one reviewer states that *tuamaneng* ‘self’ can be said, where *man* is an emphatic, which is an additional piece of evidence for the polymorphemic nature of *tuaeng*. In contrast, neither *ziji* nor *jibun* can be split in the same way as *tuaeng*.

⁶ One reviewer pointed out that *tuaeng* may actually encode some phi features because, for some native Thai speakers, it cannot take an inanimate entity as its antecedent.

⁷ Native Thai speakers’ judgments seem to vary on whether *tuaeng* can be bound by the long-distance subject *John* in (22c). In fact, Deen and Timyam’s (2018) experimental study showed that native Thai speakers prefer the local binding of *tuaeng*. I leave this question open.

Third, the three anaphors are subject-oriented (Aikawa 2002; Haddad 2007; Huang et al. 2009). As shown in (23a)-(23c), they can be bound only by the subject NP *John*.

(23a) *John_i yijing tongzhi Bill_j ziji_{i/*j}-de fenshu le.* (Chinese)
 John already inform Bill self-GEN grade PST
 ‘John_i already told Bill_j his_{i/*j} grade.’ (Huang et al. 2009:337)

(23b) *John_i-ga Bill_j-ni jibun_{i/*j}-nitsuite hanashi-ta.* (Japanese)
 John-NOM Bill-DAT self-about tell-PST
 ‘John_i told Bill_j about himself_{i/*j}.’ (Aikawa 2002:157)

(23c) *John_i bok Bill_j wa tuaeng_{i/*j} cha tong pai* (Thai)
 John tell Bill that self will need go
 ‘John_i told Bill_j that he_{i/*j} will need to go.’ (Haddad 2007:369 (26) modified)

Fourth, all three anaphors can be the head of RCs.

(24a) *John bu xihuan yizhi dui ren lengmo de ziji.* (Chinese)
 John not like always to people cold DE self
 ‘John does not like himself who is always cold to people.’

(24b) *John_i-wa Mary_j-ni tsumetaku atat-ta jibun_i-o seme-ta.* (Japanese)
 John-TOP Mary-DAT cold treat-PST self-ACC blame-PST
 ‘John blamed himself who was hard on Mary.’ (Aikawa 2002:158)

(24c) *chan mai chop tuaeng thi mai mi kwamsuk.* (Thai)
 I not like self that not have happiness
 ‘I do not like myself who is not happy.’

Thus, *Tuaeng*, *ziji* and *jibun* share at least four properties that are lacking in other languages such as English. Considering these similarities, this paper assumes that the three anaphors resemble each other, and they are expected to work similarly within RCs.

5.2 Participants

A total of 14 native Thai speakers participated in this experiment. They were graduate students from a university in Thailand and a university in the US. Their ages ranged from 23 to 30, and none of them have experience of living outside Thailand before the age of 18. They were paid to participate in this experiment.

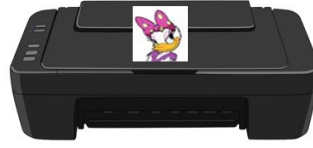
5.3 Materials and design of the experiment

Four Disney characters, *Mickey*, *Minnie*, *Donald*, and *Daisy*, were used in the task. They were introduced at the beginning of the experiment, followed by four multiple-choice questions to confirm that the participants were familiar with the names of the characters. The participants were told that the Disney characters always put their face photos on their belongings. For example, if a picture shows Mickey’s face photo on a printer, as in (25a), it means the printer belongs to Mickey. The experimental sentence for that picture is shown in (26).

(25a)



(25b)



- (26) *Daisy thotsop khrueangphim khong tuaeng thi Mickey som*
 Daisy test printer GEN self that Mickey repair
 ‘Daisy tested self’s printer that Mickey repaired.’

For each experimental stimulus, the participants were presented with a single picture such as (25a) or (25b) and a sentence such as (26) at the same time. They were asked to judge whether the sentence and the picture matched by selecting one of two choices: *sot khlong* ‘Match’ or *mai sot khlong* ‘Mismatch’. Importantly, in order for the sentence (26) to match the picture (25a), *khrueangphim khong tuaeng* ‘self’s printer’ must be interpreted as Mickey’s printer. In other words, the antecedent of the anaphor *tuaeng* must be the RC subject *Mickey*.

This experiment involves one factor (*Antecedent Position*), which results in two critical conditions: i. Matrix Subject (the anaphor *tuaeng* is intended to be co-indexed with the matrix subject) and ii. RC Subject (the anaphor *tuaeng* is intended to be co-indexed with the RC subject). Consider (26) as an example again. The two conditions based on the sample stimulus in (26) are: the Matrix Subject condition (25b) and the RC Subject condition (25a). The pictures vary in their display of the printer’s ownership (*Daisy/Mickey*) in the two conditions. On the one hand, in order for (26) to be judged as a true statement with (25b), the anaphor must be able to refer to the matrix subject *Daisy*. On the other hand, in order for (26) to be judged as a true statement with (25a), the anaphor must be able to refer to the RC subject *Mickey*.

For each condition, a total of 24 sentences with different lexicalizations were created. Each of the 24 sentences was then combined with one picture where the anaphor is intended to refer to the matrix subject and another picture where the anaphor is intended to refer to the RC subject, resulting in a total of 48 sentence-picture pairs⁸. These 48 pairs were distributed into two lists so that there were 24 critical items in each list, which contained only one condition from the same lexicalization. Each list has 12 items of the same condition.

Moreover, a set of fillers (Type 1 and Type 2 fillers) were included: the Type 1 fillers involved 24 items and the Type 2 fillers involved 12 items. First, each Type 1 filler has a ditransitive verb and there is always an anaphor embedded inside the direct object NP. One sample item is shown in (27).

- (27) *Donald hai photchananukrom khong tuaeng kae Minnie.*
 Donald give dictionary GEN self to Minnie
 ‘Donald gave self’s dictionary to Minnie.’

The pictures for the two conditions are:

(28a)



(28b)



Since *tuaeng* is subject-oriented, in (27), it can refer only to the subject NP *Donald*, not the dative NP *Minnie*. The two conditions for each Type 1 filler are a matching condition in which a picture is such that the

⁸ In order to rule out the potential confounding factors of *gender* (i.e., male vs female) and *animal type* (i.e., mouse vs duck) of the characters, *Mickey* was always paired with *Daisy* and *Minnie* was always paired with *Donald* in the stimuli.

anaphor is intended to be co-indexed with the subject NP (Type I-Subject) and a mismatching condition in which a picture is such that the anaphor is intended to be co-indexed with the dative NP (Type I-Dative). The Type 1 fillers were used to monitor (i) whether participants paid enough attention to the experimental items and (ii) whether they had the expected subject-orientation for *tuaeng*. Based on the binomial distribution, out of 12 items, participants were expected to accept 9 items or more in Type I-Subject and reject 9 items or more in Type I-Dative. The results showed that one participant wrongly accepted 10 items in Type I-Dative and her data were removed.

In addition, there were 12 Type 2 fillers. For each item, the anaphor had two possible interpretations. One item with its two conditions is shown below.

(29a)



(29b)



- (30) *Minnie_j hen Donald_k khap rot khong tuaeng_{i/k}*
 Minnie see Doanld drive car GEN self
 ‘Minnie_j saw that Donald_k drive self_{i/k}’s car.’

In (30), the anaphor *tuaeng* can refer to either the matrix subject *Minnie* or the embedded subject *Donald*. There were two conditions for the Type 2 fillers: (i) one condition in which a picture was such that the anaphor was intended to refer to the matrix subject (Type II-Matrix) and (ii) one condition in which the picture was such that the anaphor was intended to refer to the embedded subject (Type II-Embedded). There were 6 items for each condition. Since the anaphor can refer to either the matrix subject or the embedded subject, participants were expected to consistently accept the items in both conditions. The Type 2 fillers were used to monitor whether there were any participants who used an irrelevant strategy to make judgments. If our participants consistently rejected the items of the RC Subject condition as well as the Type II-Embedded condition, they might just choose to reject any item where the anaphor is intended to refer to an embedded subject. The results showed that all participants accepted at least 4 out of 6 items in Type II-Matrix. Based on this, we can infer that they did not simply reject the items where the anaphor is intended to refer to the embedded subject.

Procedure

The experiment was conducted on an online survey website. Although there was no time limit, all participants were able to finish the experiment within 20 minutes. Before seeing the actual experimental items, the participants were presented with two examples, which were used to acquaint them with the picture-matching truth-value judgment experiment. The sentences in the two examples were structurally different from those in the critical items and fillers. After that, the participants continued to look at a set of examples, which were created to help them understand the rule that they ought to select ‘Match’ as long as there is one possible interpretation of the given sentence that matches the given picture. This is crucial because participants may reject an interpretation that is acceptable but less preferable (White, Bruhn-Garavito, Kawasaki, Pater, & Prévost 1997). There were three examples, all of which share the same sentence in (31).

- (31) *Mickey_i bok Daisy_j wa Donald_k chai yasifan khong tuaeng_{i*/j/k} motlaeo.*
 Mickey tell Daisy that Donald use toothpaste GEN self finished
 ‘Mickey_i told Daisy_j that Donald_k already used up self_{i*/j/k}’s toothpaste.’



Three Disney characters were used in this set of example items. In (31), the anaphor *tuaeng* can be co-indexed with either the matrix subject *Mickey* or the RC subject *Donald*, but not *Daisy*, due to its subject-oriented property. Participants saw the three pictures (32a), (32b) and (32c) consecutively, each of which was combined with (31). First, (32a) was presented and the participants clicked ‘Match’ after understanding that *yasifan khong tuaeng* ‘self’s toothpaste’ can be interpreted as *Mickey’s toothpaste*. Next, (32b) was presented and the participants clicked ‘Mismatch’ after understanding that *yasifan khong tuaeng* ‘self’s toothpaste’ cannot be *Daisy’s toothpaste*. Last, (32c) was presented and the participants clicked ‘Match’ after understanding that *yasifan khong tuaeng* ‘self’s toothpaste’ can be *Donald’s toothpaste*. After going through the examples, participants continued to practice four trials before the actual experimental items.

Findings

Table 1 summarizes the participants’ mean proportions of ‘Match’ answers, standard deviations (SDs), and standard errors (SEs) of the two critical conditions.

Table 1. Participants’ mean proportions of ‘Match’ answers of the critical conditions in Thai

Subject Position	Mean (SD)	SE
Matrix Subject	0.97 (0.07)	0.02
RC Subject	0.05 (0.09)	0.03

The data clearly shows that the participants prefer the co-reference between the matrix subject and the anaphor over the co-reference between the RC subject and the anaphor. Pairwise comparison showed a significant difference between the two conditions ($t(12) = 21.12, p < .01$).

Now we compare our Thai data with Chen’s (2019) Chinese and Japanese data. Table 2 summarizes the 69 native Chinese participants’ and 28 native Japanese participants’ judgments of the two critical conditions in Chinese and Japanese respectively:

Table 2. Participants’ mean proportions of ‘Match’ answers of the critical conditions in Chinese and Japanese (Chen 2019)

Language	Subject Position	Mean (SD)	SE
Chinese	Matrix Subject	0.83 (0.19)	0.02
	RC Subject	0.85 (0.21)	0.03
Japanese	Matrix Subject	0.94 (0.09)	0.02
	RC Subject	0.1(0.13)	0.02

Comparing the Thai data in Table 1 and the Chinese and Japanese data in Table 2, we can see that Thai and Japanese RCs behave similarly: in both languages, the anaphor within the head NP can only refer to the matrix subject. In contrast, the interpretation of the anaphor is more flexible in Chinese: it can refer to either the matrix subject or the RC subject. Since previous studies have claimed that the head NP of RCs is base-generated in Japanese but raised in Chinese (e.g., Aoun & Li 2003; Fukui & Takano 2000), the above comparison strongly suggests that Thai and Japanese share the same base-generation strategy to derive the head NP for RCs.⁹

⁹ In the fillers of the Thai experiment, the anaphor *tuaeng* always takes as its antecedent an element that is structurally higher while in the critical items, *tuaeng* is intended to refer to an element that is either higher or lower in the structure. One may argue that this creates a bias towards ‘Mismatch’ answers in the RC Subject condition, since there is a priming effect from the fillers for upward orientation. However, Chen’s (2019) Chinese result suggests

Also, the individual participants' judgments were examined. As mentioned earlier, since there were 12 items in each condition, based on the binomial distribution, if a participant accepted or rejected 9 or more items, we would be 95% confident that her judgments were consistent. The results reveal that all 13 participants accepted 9 items or more in the Matrix Subject condition and rejected 9 items or more in the RC Subject condition, which indicates that they consistently accepted the co-reference between the matrix subject and the anaphor but consistently rejected the co-reference between the RC subject and the anaphor. The individual data were in line with the group data, implying that in Thai RCs, the anaphor inside the head NP cannot refer to the RC subject.

Discussion

The group and individual data of the truth value judgment experiment suggest that the anaphor *tuaeng* 'self' inside the head NP of Thai RCs cannot take the RC subject as its antecedent, which supports the head external analysis for Thai RCs. If the head is base-generated external to the RC, what is the empty category inside the RC? While Hoonchamlong (1991) claimed that Thai RCs do not involve any type of movement, Jenks (2011) conducted an online experiment with native Thai speakers and showed that relativization is subject to island constraints (Ross 1967), as in (33a), (33b) and (33c).

(33a)* *wanni chan hen [ma_i [thi Nit ruchak dek_j [thi ec_i kat ec_j]]]*
 today I see dog that Nit know child that bite
 'Today I saw the dog that Nit knows the child that bit.'

(33b)* *wanni chan hen [ma_i [thi Nit ruchak dek_j [thi ec_j ying ec_i]]]*
 today I see dog that Nit know child that shoot
 'Today I saw the dog that Nit knows the child that shot.'

(33c)* *wanni chan hen ma_i thi Nit kangwon phro ec_i kat Noy*
 today I see dog that Nit worry because bite Noy
 'Today I saw the dog that Nit is worried because bit Noy.' (Jenks 2011:147)

The ungrammatical status of the above sentences indicates that the head NP cannot be extracted from a complex NP island or an adjunct island in Thai, which further implies that there is movement of some kind involved in Thai RCs. Given the finding that the head NP is base-generated external to the RC, there should be a null operator moving to [Spec, CP], as illustrated in the following example.

(34) [NP [NP *rupphap*]_i [CP Op_i *thi* [TP *Daisy at* *t_i*]]]
 photo that Daisy compress
 'the self's photo that Daisy compressed'

This is in line with the traditional head-external analysis, illustrated in (2) and repeated below.

(35) [DP the cake_i [CP op_i/which_i John ate *t_i*]]

However, the RC is right-adjoined to the head NP in (35), which is not compatible with Kayne's (1994) Antisymmetry Theory, and how the head NP is merged with the RC still needs further examination. In addition, there is another piece of evidence suggesting that the head NP is base-generated. Bianchi (1999) proposed that scope interaction can be used to test the head derivation of RCs.

that the consistent rejection of the RC Subject condition in the Thai experiment is not purely because of the priming effect. Since in all fillers of the Thai, Chinese and Japanese experiments, the anaphor takes a structurally higher element as its antecedent, we would expect comparable results for all critical items across the three languages, if there is a significant priming effect from the fillers. However, the consistent acceptance of the RC subject condition in the Chinese experiment suggests that participants are able to co-index an anaphor with a structurally lower antecedent, even if there is a potential priming effect from the fillers.

- (36) I phoned the [two patients]_i [that every doctor will examine t_i tomorrow].
 a. 2 patients > every doctor b. every doctor > 2 patients

In (36), there are two possible interpretations: (i) ‘I phoned the two specific patients that will be examined by every doctor tomorrow,’ as indicated in (a) and (ii) ‘I phoned different groups of two patients that every doctor will examine tomorrow,’ as indicated in (b). The existence of the interpretation in (b) suggests that the universal quantifier *every* can have scope over the head *two patients*. Since *two patients* c-commands *every doctor* on the surface, it implicates the raising of *two patients* from within the RC, which can reconstruct into its base position at LF. In contrast, only the interpretation in (a) is available in the equivalent Thai sentence.

- (37) *chan thoha phupuai songkhon thi mo thukkhon cha truatsop phrungni*
 I call patients two-CL that doctor every will examine tomorrow
 ‘I called the two patients that every doctor will examine tomorrow.’
 a. 2 patients > every doctor b. *every doctor > 2 patients

Thus, the contrast between Thai and English in terms of the scope interaction in RCs further supports the head base-generation analysis for Thai RCs.

6 Conclusion

There have been several widely used diagnostics to examine the head derivation of RCs. Jenks (2011) used idioms and adjectival modifiers to argue that the head NP of Thai RCs should be raised from within the RC. However, a comparison between Thai and Chinese in terms of the reconstruction effects of their equivalent idioms has shown that the head NP in Thai and Chinese RCs may be subject to different derivational mechanisms. Moreover, using adjectival modifiers as a diagnostic is problematic (Heycock 2005).

This paper focuses on anaphor binding, another well-established diagnostic, to examine the head derivation of Thai RCs. A controlled truth value judgment experiment was created, and the experimental results suggest that the anaphor *tuaeng* ‘self’ inside the head NP of Thai RCs cannot be co-indexed with the RC subject. By comparing the Thai data with Chen’s (2019) Japanese and Chinese data, we can see that Thai is parallel to Japanese but is in a clear contrast to Chinese. Since the head NP is base-generated in Japanese but raised in Chinese (e.g., Aoun & Li 2003; Fukui & Nakano 2000), this comparison further argues for the head base-generation analysis for Thai RCs. However, how the head NP is merged with the RC still needs further investigation.

References

- Aikawa, Takako. 2002. Reflexives. In N. Tsujimura (Ed.), *Handbook of Japanese Linguistics*, 154-190. Blackwell Publishing.
- Aoun, Joseph & Yen-hui Audrey Li. 2003. *Essays on the representational and derivational nature of grammar: the diversity of wh-constructions*. Cambridge, MA: MIT Press.
- Bhatt, Rajesh. 2002. The raising analysis of relative clauses: Evidence from adjectival modification. *Natural Language Semantics* 10:43–90.
- Bhatt, Rajesh & Sabine Iatridou. 2012. Comments on Reconstructing Reconstruction. Paper presented at the Workshop on Resumptive Pronouns, Hebrew University of Jerusalem.
- Bianchi, Valentina. 1999. *Consequences of antisymmetry: Headed relative clauses*. New York: Mouton de Gruyter.
- Brame, Michael K. 1968. A new analysis of the relative clause: evidence for an interpretive theory, unpublished manuscript, MIT.
- Chen, Yunchuan. 2019. The acquisition of Japanese relative clauses by L1 Chinese learners. Ph.D. dissertation, University of Hawaii at Manoa.
- Chomsky, Noam. 1977. On *wh*-movement. In P. Culicover, T. Wasow, & A. Akmajian (Eds.), *Formal Syntax*, 71-132. New York: Academic Press.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, Noam. 1986. *Knowledge of Language*. NY: Praeger.

- Chomsky, Noam. 1993. A Minimalist Program for linguistic theory. In K. Hale & S. J. Keyser (Eds.), *The view from Building 20*:1-52. Cambridge, MA: MIT Press.
- Crain, Stephen & Rosalind Thornton. 1998. *Investigations in Universal Grammar: a guide to experiments on the acquisition of syntax*, Cambridge, MA: MIT Press.
- Deen, Kamil U. & Napasri Timyam. 2018. Condition C in adult and child Thai. *Language* 94.1:157-190.
- Fukui, Naoki & Yuji Takano. 2000. Nominal structure: An extension of the Symmetry Principle. In P. Svenonius (Ed), *The derivation of VO and OV*, 219-254. Amsterdam: John Benjamins.
- Haddad, Youssef A. 2007. Subject Anaphors: Exempt or Not Exempt? *Linguistic Inquiry* 38.2:363-372.
- Han, Chung-hye. 2013. On the syntax of relative clauses in Korean. *The Canadian Journal of Linguistics* 58.2:319-347.
- Heycock, Caroline. 2005. On the interaction of adjectival modifiers and relative clauses. *Natural Language Semantics* 13.4:359-382.
- Hoonchamlong. Yuphaphann. 1991. Some issues in Thai anaphora: A government and binding approach. Ph.D. dissertation, University of Wisconsin-Madison.
- Huang, C.-T. James, Yen-hui Audrey Li & Yafei Li. 2009. *The syntax of Chinese*. Cambridge, MA: Cambridge University Press.
- Hulsey, Sarah & Uli Sauerland. 2006. Sorting out relative clauses. *Natural Language Semantics* 14:111-137.
- Jackendoff, Ray. 1977. *\bar{X} Syntax: a study of Phrase Structure*. Cambridge, MA: MIT Press.
- Jenks, Peter. 2011. The Hidden Structure of Thai Noun Phrases. Ph.D. dissertation, Harvard University.
- Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
- Lasnik, Howard. 1989. *Essays on anaphora*. Dordrecht: Kluwer.
- Miyamoto, Yoichi. 2017. Relative Clauses. In M. Shibatani, S. Miyagawa, & H. Noda (Eds.), *Handbook of Japanese Syntax*, 611-634. De Gruyter Mouton.
- Murasugi, Keiko. 2000. Antisymmetry analysis of Japanese relative clauses. In A. Alexiadou:Law, A. Meinunger, & C. Wilder (Eds.), *The syntax of relative clauses*, 231-264. Amsterdam: John Benjamins.
- Ross, John. 1967. Constraints on variables in syntax. Ph.D. dissertation, MIT.
- Sauerland, Uli. 1998. The meaning of chains. Ph.D. dissertation, MIT.
- Schachter, Paul. 1973. Focus and relativization. *Language* 49:19-46.
- White, Lydia, Joyce Bruhn-Garavito, Takako Kawasaki, Joe Pater & Philippe Prévost. 1997. The researcher gave the subject a test about himself: problems of ambiguity and preference in the investigation of reflexive binding. *Language Learning* 47:145-72.

Reviewed: Received 6 August 2020, revised text accepted 19 November 2020, published 1 February 2021

Editors: Editor-In-Chief Dr Mark Alves | Managing Eds. Dr Paul Sidwell, Dr Nathan Hill, Dr Sigrid Lew