Nominal Modification in Chinese and Thai

Shi-Zhe Huang and Peter Jenks
Haverford College and University of California, Berkeley

1. Introduction

Mandarin Chinese and Thai are Generalized Classifier Languages (Chierchia 1998), allowing definite bare nouns in argument positions (1a, 2a) and requiring numeral classifiers (1b, 2b).

(1) MANDARIN CHINESE
   a. xuesheng chi-le fan
      student eat-PRF rice
      ‘The student(s) have eaten.’
   b. san *(ge) xuesheng
      three CLF student
      ‘three students’

(2) THAI
   a. nakrian kh ə y kin khaaw
      student PRF eat rice
      ‘The student(s) have eaten.’
   b. nakrian saam *(khon)
      student three CLF
      ‘three students’

Yet Mandarin Chinese and Thai have striking differences in the marking of modifiers. While Mandarin makes use of a single polyfunctional marker de to mark relative (3a), possessive (3b), and mensural modifiers (3c), Thai makes use of three different morphemes in the three different constructions (4a-c), all historically derived from nouns, to mark the three different patterns.

(3) MANDARIN CHINESE
   a. [RelativeCP]-de N
   b. [PossDP ]-de N
   c. [MeasureP]-de N

(4) THAI
   a. N thii-[RelativeCP]  
   b. N khoɔy-[PossDP]  
   c. N khanaat-[MeasureP]


In this paper we adopt the analysis of de as a type-shifter (i.e., <<e,t>,e>, proposed by Huang (2006) to account for the distribution of simple and complex adjectives in modificational environments, an analysis which we review in Section 2. In Section 3 we demonstrate that this analysis cannot be extended to the Thai markers of modification, which we show are predicate-forming operators; we conclude that nominalization performed by de is done covertly in Thai. In Section 4 we demonstrate how dialectal data reveals that de might be the fusion of two distinct functional markers (Zhu 1961, 1993, Huang 2006), one for predication/abstraction, just like the Thai modifier markers, and one for nominalization, as Huang (2006) claimed. We show in section 5 that by unpacking de into two distinct functions, namely predication and nominalization, and by positing a silent nominalizer in Thai modification structure, adopting the Type-Matching Constraint for Thai, some apparent differences between Chinese and Thai modifier markers in their ability to license NP-ellipsis (e.g. Simpson 2003, Saito et al 2008, ans Cheng & Sybesma 2009) can be explained. We present some challenges at the end of the paper.

2. Chinese de as a type-shifter (Huang 2006)

The original motivation for treating de as a type shifter of the sort <<e,t>,e> arises out of the following paradigm as noted in the classic work of Zhu Dexi (e.g. 1956, 1961):
a. na ge diqu *(hen) pinqiong
that CLF region very poor
‘That region is very poor.’

b. pinqiong (de) diqu
poor de region
‘the/a poor region’

c. hen pinqiong *(de) diqu
very poor de region
‘the/a very diligent student’

In order to explain the complementary distribution of bare adjectives and complex adjectives in the predicate position as well as in the modifier position, Huang (2006) proposes a type-theoretical account on modification which captures a generalization based on the data:

(6) TYPE-MATCHING CONSTRAINT ON MODIFICATION (Huang 2006)
A bare noun and its modifier must be of the same semantic type.

Crucially, Huang (2006) adopts the view that nouns in Mandarin Chinese denote kinds, of type $e$ (Krifka 1995, Chierchia 1998), and also argues that adjectives in Mandarin Chinese denote nominalized properties, also of type $e$, cf. women yao zhangsheng pinqiong ‘We should wipe out poverty.’ (Huang 2006, p. 350). *Hen*, which is type $<e, <e,t>>$, the “up” operator of Chierchia (1998), converts nominalized property to a predicate, so that [hen SA] can serve as predicates (5a), while *de*, which is type $<<e,t>, e>$, the “down” operator of Chierchia (1998), nominalizes a predicate to type $e$ (5c), so as to convert a predicate modifier to be matched with the modificiee, which is type $e$, satisfying the Type-Matching Constraint on modification.2

Another classifier language, Thai, shows none of these asymmetries in the distribution of bare and complex adjectives, as both can serve either as sentence predicates or nominal modifiers:

(7) a. Nat suuŋ *(maak)
Nat tall very
‘Nat is (very) tall.’

b. Nat ruu caak nakrian suuŋ maak
Nat knows of student tall very
‘Nat knows the/a very tall student.’

And unlike Chinese, Thai adjectives do not seem to be inherently nominalized, as they can occur with an overt derivational nominalizer, *khwaam* (Prasithrathsint 2005):

(8) *(khwaam)-khayan pen siŋ thii dii
QUALITY-diligent COP:PRED thing REL good
‘Diligence is a good thing.’

Typological differences aside, it is still unclear why the Type-Matching Constraint does not seem to hold for examples such as (7b); we return to this issue in section 4.

3. Thai modification and predicate-forming nominal heads

The polyfunctional marker of modification in Chinese *de* occurs with relative clauses, possessives, and attributive measure phrases, as well as adjectives, as the previous section showed:

---

1 We assume non-contrastive intonation and a matrix clause structure. Other sentence types and intonational variations might make bare adjectival predicates possible. (Huang and Li 2009, Liu 2010 a.o.)

2 Huang (2006) argues that the two type $e$ entities compose through simultaneous type lifting.
Yet these three types of modifiers are marked with three different particles in Thai:

(12) nakrian *(thii) chan chɔɔp

student REL 1SG like
‘the/a student that I like’

(13) napsit *(khɔɔŋ) chan

book POSS 1SG
‘my book’

(14) thurian *(khanaat) saam loo

durian SIZE 3 kilo
‘a/the three-kilo orange’

The Thai markers of modification in (12)-(14) are all synchronically related to nouns. *Thii* in (12) is synonymous with a noun meaning ‘place,’ the marker of possession *khɔɔŋ* in (13) is synonymous with a noun meaning ‘possession,’ while *khanaat* is synonymous with a noun meaning ‘size.’

The nominal nature of these words initially seems to support an analysis of these Thai markers along the lines of Huang (2006)’s nominalization analysis of *de*; after all, the Thai words are nouns. However, we will show in this section that the Thai particles do not type-shift properties to nominalized properties at all, but rather serve to form predicates. This indicates that the Thai markers of modification should not be analyzed as nominalizing type-shifters.

3.1. Relative clauses

One interesting property of the marker *de* in Chinese relative clauses is that it is obligatory in both subject and object relative clauses:

(15) a. wo xihuan ___ * *(de) gexing

1SG like DE singer
‘the/a singer that I like’

b. ___ i xihuan wo *(de) gexing

like 1SG singer
‘the/a singer(s) that like(s) me’

In contrast, Thai *thii* is optional with subject relative clauses (Kuno & Wongkhomthong 1981b):

(16) a. nakrian, *(thii) chan chɔɔp ___

student REL 1SG like
‘the/a student(s) that I like’

b. nakrian, *(thii) ___ i chɔɔp chan

student REL like 1SG
‘the/a student(s) that like(s) me’

Kuno & Wongkhomthong (1981b) show that when *thii* is omitted, relative clauses have generic rather than specific interpretations.

Jenks (to appear) argues that *thii* is a relative operator in C⁰. This provides a natural explanation for why *thii* is optional with subject relative clauses: *thii*-less relatives are structurally reduced participles that lack an external argument. This provides a natural account for the genericity of these kinds of relative clauses:

(17) a. [CP thiix [TP ___ [VP chɔɔp chan ]]] ⇒ λx[like(SPEAKER,x)]

b. [VP chɔɔp chan ] ⇒ λx[like(SPEAKER,x)]
In (17a), a full relative clause with relative operator intact is interpreted as a property abstracted over the subject, with saturated world and time variables. In contrast, (17b) represents the reduced relative clause. The subject positions is never saturated, so it never needs to be abstracted over; Jenks assumes that the world and time variables are simply absent. This accounts for the fact that these participial relatives lack temporal specification. Thus, an analysis of thii as a semantically contentful relative operator provides a natural explanation for why it can only be absent with subject relative clauses.

Further evidence that thii is an operator comes from free relative clauses in Thai, shown below:

(18) a. (sin) thii khun hen khii tik ‘Sears’
    thing REL you see COP:EQ building
    ‘What you see in front of you is the Sears Tower.’

b. chan may chia (sin) thii khun book
    I not believe thing REL you say
    ‘I don’t believe what you said.’ (Hoonchamlong 1991, p. 180-181)

While the evidence is circumstantial, it is nevertheless true that free relatives are only licensed when headed by an overt operator in English; likewise, Thai free-relatives must occur with the operator thii. These free relative clauses can be preceded by a ‘dummy’ noun meaning ‘thing’, but these are not elliptical contexts. Section 5 below illustrates that Thai does not allow N-ellipsis at all.

Together, then, the optionality of thii with subject relatives and the existence of free relative clauses support an analysis of thii as a relative operator (see Jenks to appear for a more detailed analysis along these lines).

3.2. Possessive DPs

As we saw at the beginning of this section, where Chinese uses de to mark possessive noun phrases, Thai uses the word khɔɔŋ before the possessor, a word which is otherwise a noun meaning ‘possession.’ This nominal khɔɔŋ can also occur as the main predicate of a clause in Thai:

(19) a. nãnsii lem nii khɔɔŋ chan
    book CLF this POSS I
    ‘This book belongs to me.’

b. nakrian khon nii khɔɔŋ chan
    student CLF this POSS I
    ‘This student belongs to me.’

The fact that possessives marked with khɔɔŋ might lead to the hypothesis that khɔɔŋ has a verbal use as well, especially as PP and nominal predicates in Thai are obligatorily preceded by copula.

Yet khɔɔŋ in (19) is not verbal. Evidence comes from negation, which only occurs before verbs:

(20) a. *nãnsii lem nii maj khɔɔŋ chan
    book CLF this not POSS I
    ‘This book is mine.’

b. nãnsii lem nii maj chaj khɔɔŋ chan
    book CLF this not COP:NPI POSS I
    ‘This book doesn’t belong to me.’

In (20b), the possessive copula is possible, but negation must precede the copula chaj which occurs in negative contexts (Chiravate 1999). The positive counterpart of (20b) is in (21), where the possessive phrases marked with khɔɔŋ occurs as the complement of the predicative copula pen, which only occurs with nominal predicates (Kuno & Wongkhomthong 1981a):

(21) nãnsii lem nii pen khɔɔŋ chan
    book CLF this COP:PRED POSS I
    ‘This book is mine.’
There is an interesting restriction on these postcopular possessive predicates, however; they can only serve as predicates for inanimate possessees:

(22) *nakriian khon nii pen khɔɔŋ chan

student CLF this COP:PRED POSS I

(intended) ‘This student is mine.’

One way of thinking about this distinction is that in (21), the sentence has the literal meaning ‘the book is my possession,’ whereas that interpretation is unavailable in (22). This means that in predicative contexts, the lexical meaning of the noun khɔɔŋ is preserved; this must not be true in adnominal possessives (cf. 19b).

To summarize, then, the predicative distribution of [khɔɔŋ XP] indicates that the phrase [khɔɔŋ] is a predicate, characterizing the property of being a possessum of XP. As the predicative copula pen only can occur with nominal complements, we can conclude that khɔɔŋ ‘possession, belonging’ is still a noun, a fact which finds support from its inability to be negated.

So khɔɔŋ must have a denotation that takes an individual argument and renders a property of being possessed by that individual. In other words, it is the two-place nominal predicate ‘possession of’:

(23)

a. [[khɔɔŋ chan]]

b. = λxλy [possession-of(x,y)]([chan])

c. = λxλy [possession-of(x,y)](SPEAKER)

d. = λy [possession-of(SPEAKER,y)]

So the denotation of the [khɔɔŋ XP] possessor is the set of elements that ‘belong’ to XP.

Another argument that khɔɔŋ is a predicate forming operator is that it is optional with kinship terms (24b). This is also true for de in Chinese, though only with pronominal and proper noun possessors (24a):³

(24)

a. wo (de) mama

1sg de mother

‘my mother’

b. mee *(khɔɔŋ) chan

mother POSS I

‘my mother’

Because such kinship terms are relational nouns, they can directly take the possessor as an argument, obviating the need for the possessive particle:

(25)

a. [[wo mama]]

b. = λxλy [mother-of(x,y)]([wo])

c. = λxλy [mother-of(x,y)](SPEAKER)

d. = λy [mother-of(SPEAKER,y)]

While the derivation in (25) is straightforward, it is not exactly clear how this interpretation is achieved when the possessive markers are present with relational nouns.

In summary, we have seen evidence that possessive modifiers occurring after khɔɔŋ are predicative, and that khɔɔŋ is a noun. These facts support the idea that khɔɔŋ is a transitive nominal which takes the possessor as an argument, forming a possessive predicate which can then modify the head noun. It is not clear exactly how Chinese de fits into this picture, but it is notable that the distribution of de is optional with relational nouns khɔɔŋ.

3.3. Attributive measure phrases

The third case where Chinese uses the polyfunctional modification marker de is with attributive measure phrases (Jiang 2009, Li & Rothstein 2012). Attributive measure phrases in Thai are introduced by the modification marker khanaat ‘size’:

³ Thai khɔɔŋ is also generally optional with ‘typically possessed objects’ (‘book’ but not ‘tree’); this may be a pragmatic effect.
Attributive measure phrases can be marked with *khanaat* regardless of whether they actually indicate a size; (26a) would most typically the width of a rope, while (26b) describes volume. Without *khanaat*, true (non-attributive) measure readings emerge:

\[(27)\]
\[
\begin{align*}
a. \text{saay saam sen} & \quad \text{b. \(\text{khuat-}\)naam saam lit} \\
\text{rope size three cm.} & \quad \text{bottle-water size three liter} \\
\text{‘the/a three-centimeter rope’} & \quad \text{‘the/a three-liter bottle of water’}
\end{align*}
\]

The contrast between (26) and (27) illustrate that *khanaat* is responsible for the attributive readings of measure phrases.

Attributive measure phrases can also be marked with dimensional adjectives:

\[(28)\]
\[
\begin{align*}
a. \text{\(\text{s}i\text{ɨa}\) yaaw saam met} & \quad \text{b. \text{thurian nak saam loo}} \\
\text{rope long three cm.} & \quad \text{durian heavy three kilo} \\
\text{‘three-yard long rope’} & \quad \text{‘three-kilo-heavy durian’}
\end{align*}
\]

The examples above might suggest that *khanaat* might be a kind of adjective. However, the dimensional adjectives in (28) can co-occur with *khanaat*, although somewhat unnaturally:

\[(29)\]
\[
\begin{align*}
a. \text{\text{\text{\(\text{s}i\text{ɨa}\) khanaat yaaw saam met}} & \quad \text{b. \text{\(\text{th}i\text{ɨ}n\) khanaat nak saam loo}} \\
\text{rope size long three cm.} & \quad \text{durian size heavy three kilo} \\
\text{‘three-yard long-sized rope’} & \quad \text{‘three-kilo-heavy-sized durian’}
\end{align*}
\]

We propose that while the dimensional adjectives in (28) introduce the scale which is measured by the measure phrase, the noun *khanaat* serves to convert this interval into a property which can modify the noun. Like the relative operator *thii* and the possessive marker *\(\text{kh}\text{ɔɔŋ}\)*, *khanaat* is a noun.

Evidence that the measure phrase which includes *khanaat* is a predicate comes from the ability of *khanaat* phrases to function as main predicates. This option is also available for the dimensional adjectives with measure phrases in (28):

\[(30)\]
\[
\begin{align*}
a. \text{saay \text{s}n\text{ɨ ni\text{n}} khanaat saam sen} & \quad \text{b. saay \text{s}n\text{ɨ ni\text{n}} yaaw saam met} \\
\text{rope CLF this size three cm} & \quad \text{rope CLF this long three meter} \\
\text{‘This rope is a three-centimeter one.’} & \quad \text{‘This rope is three meters long.’}
\end{align*}
\]

Negation again shows that *khanaat* is not a verb, however (31a). Negation of a measure phrase must involve the introduction of a verb such as *\text{tíy}*, which can in turn be negated:

\[(31)\]
\[
\begin{align*}
a. \text{\*saay \text{s}n\text{ɨ ni\text{n}} maj khanaat saam sen} \\
\text{rope CLF this not size three cm} \\
\text{‘This rope doesn’t reach three-centimeters.’}
\end{align*}
\]

The ability of *khanaat* to occur as a main predicate (30a) along with the inability of *khanaat* to be negated indicates that, like *thii* and *\(\text{kh}\text{ɔɔŋ}\)*, *khanaat* in these contexts is a nominal which forms attributive predicates.

*Khanaat* and the measure adjectives in (24) receive a similar analysis to the possessive *\(\text{kh}\text{ɔɔŋ}\)* in the previous section: *khanaat* is a two-place relation between a measure and an individual:
(32)  a. \([\text{khanaat saam sen}])_{g,w}
    b. \(= \lambda \mu \lambda y[\text{size}(x,y) \& x = \mu]([\text{saam sen}])
    c. \(= \lambda \mu \lambda y[\text{size}(x,y) \& x = \mu](3\text{cm})
    d. \(= \lambda y[\text{size}(x,y) \& x = 3\text{cm}]

We have ignored the role of the dimensional adjective, which indicates that semantics in (32) is oversimplified. The adjective, which can be implicit, defines the scale on which the measure is set, and the degree on this scale is the actual measure argument.

In summary, we have seen that \(\text{khanaat}\) forms a natural class with \(\text{thii}\) and \(\text{khɔɔŋ}\) in functioning as a predicate. Independent arguments have been provided for the predicative status of all three modificational expressions, indicating that \(\text{thii/khoong/khanaat}\) are all predicate forming nominal heads. In conjunction with the conclusions of the previous section on Chinese, then, we have seen that the distribution of the markers of modification in Thai are strikingly different from their Chinese counterpart \(de\), which shifts the type of the modifier to match the type of the head noun.

4. The dual function of \(de\)

In this section we show that the there is more to Mandarin \(de\) than meets the eye. Zhu (1993) describes a number of Chinese dialects which, in place of Mandarin \(de\), have two separate morphemes (See Huang 2006 for a fuller discussion in English of Zhu’s data). The following examples come from Daye (Jinhu) Dialect (Wang 1991), one of the ten dialects discussed in Zhu (1993):

\[
\begin{align*}
\text{(33) a.} & \quad \text{tai jian yishang huaihuiq i ta} \\
& \text{this CLF shirt gaudy} \\
& \text{‘This shirt is gaudy.’}
\end{align*}
\]

\[
\begin{align*}
\text{b.} & \quad \text{huaihuiq i ta ko yishang} \\
& \text{gaudy shirt} \\
& \text{‘(the) gaudy shirt’}
\end{align*}
\]

The Mandarin counterpart of (33a) would have \(de\) in place of \(ta\), and the Mandarin counterpart of (33b) would have \(de\) in place of \(ta\ ko\). Based on dialectal (and historical data), Zhu proposes that Mandarin \(de\) has three distinctive functions:

- \(de1\): for adverbial modification (irrelevant for our purpose)
- \(de2\): for predicate marking, i.e., \(de2 = ta\)
- \(de3\): for nominalization i.e., \(de3 = ko\)

Representing the dialectal data schematically as: [XP \(de2\ \ de3\ N\)], Huang (2006:347, 2012), adopting Zhu’s insight, suggests that Modern Mandarin \(de\) is a fusion of \(de2\) and \(de3\), namely: \(de = de2 + de3\). Li (2012) presents similar data regarding Min (Taiwanese), where two tonally distinct heads can occur with modifiers.

Along with analysis of Thai modification in Section 3, this dialectal data suggests that while \(\text{thii/khɔɔŋ/khanaat}\) are solely responsible for predicate formation in Thai, Modern Mandarin \(de\), which is \(de2 + de3\), has the dual function of predication, which corresponds to \(\text{thii/khɔɔŋ/khanaat}\) in Thai, and nominalization, which has a silent counterpart in Thai. This conclusion provides an explanation for why Thai does not exhibit the same kind of Type-Matching Constraint violations as Chinese (compare (5c) to (7b)). If the nominalization function of \(de3\) is covert in Thai, nominalization can freely apply to the modifier in examples like (7b).\(^4\)

To summarize, these observations support a particular view of Thai and Chinese modification. Chinese \(de\) is the fusion of a predicativizing operator, always overt in Thai, and a nominalization head, which is covert in Thai.

\(^4\) Chierchia’s Blocking Principle (Chierchia 1998, p. 360) is irrelevant here since \(khwaam\) cannot be the overt counterpart in Thai of \(de3\). As stated in Section 3, \(khwaam\) is a derivational prefix, applying exclusively to adjectives and stative verbs from which it derives nouns (Prasithrathsint 2005). It is different from a functional head that type shifts without changing the categorical features of its argument.
In the following section, we show that while the structural representation in (34) clarifies the parallels and differences between Chinese and Thai, it has problems in accounting for NP-ellipsis in Chinese.

5. NP-ellipsis in Chinese and Thai

One problem for the syntactic analysis of Chinese de in (34b) is that de has been shown to license NP-ellipsis: (Saito et al. 2008, Cheng & Sybesma 2009, Li 2012):

(35) \[Wo zuotian kanjian de nanhai] bi [ni kanjian de (nanhai)] geng youqian
I yesterday see DE boy than you see DE boy more rich
‘The boy I saw yesterday is richer than the boy you saw.’ (Saito et al. 2008, ex. 56)

(36) \[wo de xuesheng] bi [ni de (xuesheng)] geng youqian
I de student than you DE student more rich
‘My student is richer than your student.’

(37) \[wu li-mi de pingguo] bi [shi li-mi de (pingguo)] haochi
5 cm. DE apple than ten cm. DE apple delicious
‘The five-cm. apple is more delicious than the 10-cm. apple.’

Example (35) illustrates that relative clause de can license the ellipsis of the nominal head in elliptical constructions such as comparatives. The same point is illustrated for possessive de in (36) and mensural de in (37).

In contrast, Thai thii/khɔɔŋ/khanaat do not license ellipsis in this same context: 5

(38) \[dek thii chan hen mueawaan] ruy kwaaw [*(dek) thii thoɔ hen]
child REL I see yesterday rich than child REL you see
‘The child I saw yesterday is richer than the child you saw.’

(39) \[nakrian khɔɔŋ chan] ruy kwaaw [*(nakrian) khɔɔŋ thoɔ]
student POSS I rich than student POSS you
‘My student is richer than your student.’

(40) \[thurian khanaat saamsip sen] aroy kwaaw [*(thurian) khanaat yiisip sen]
durian SIZE 30 cm. delicious than durian SIZE 20 cm.
‘The 30-cm. durian is more delicious than the 20-cm. durian.

In a sense, this syntactic distinction between Chinese and Thai provides further evidence that the markers of modification in the two languages should not receive the same analysis.

A common assumption about ellipsis is that it is licensed for the complement of an overt functional head with a filled specifier (Lobeck 1995). The Thai modification markers are correctly predicted to disallow N’-ellipsis under this view because the modification markers are not functional

---

5 NP-ellipsis in these examples would be licensed by the addition of a classifier between the noun and the modifier (cf. Jenks 2011:90-93). The licensing of classifiers by modifiers is discussed in Jenks 2011, ch. 5.
projections of the noun. However, under the analysis of de in (34b), it is not clear why ellipsis is allowed, as neither is de a functional projection of the noun.

Faced with similar data and considerations, Li (2012) concludes that Taiwanese equivalents of de2 and de3 form constituents with the modifier (XP) and NP, respectively: \([\text{XP } \epsilon^0][\epsilon^7 \text{NP}])\), which we render structurally for Mandarin Chinese as (41):

\[
(41)
\]

\[
\begin{array}{c}
\text{XP} \\
\text{de2} \\
\text{de3} \\
\text{NP}
\end{array}
\]

Under Lobeck’s theory of ellipsis, the ability of de3 to license the deletion of NP is expected as de3 takes the NP as its complement, and because its specifier is filled. Along similar lines to Li (2012)’s analysis, Simpson (2003) and Saito et al. (2008) argue that de is a determiner, and Cheng & Sybesma (2008) argue that de is a classifier. However, in many of these cases, the main argument for the analysis of de as a determiner or classifier is precisely the ellipsis licensing facts in (35)-(37).

Can the syntactic analysis of de2+de3 in (41) be reconciled with the composite head view of de in (34)? There are at least two obstacles to the unification of these two analyses. The first obstacle is morphological: some theories assume that a precondition for morphological fusion is the head-complement relationship (cf. Brody 2000, Williams 2003). Clearly, de3 is not the complement of de2 in (41). However, there are proposals for, e.g., English possessive pronouns which take them to be mergers of a possessor with Saxon genitive ‘s, e.g., me+’s \(\rightarrow\) my (Matushansky 2006, p. 86). If this were true, we could take the merger of de2+de3 to be a roughly comparable case of merger between a clitic and its specifier. In conclusion, this morphological obstacle does not seem too severe.

The second obstacle to bringing the analysis of Chinese in (34) in line with the structure in (41) is semantic. The problem is that de3 forms a constituent with the head noun to the exclusion of the modifier in (41), so compositionality demands that de3 compose with the head noun. However, de3 is hypothesized to serve as a nominalizing type-shifter on the modifier, so the hypothesized semantic argument of de3 is the predicate headed by de2, with which de3 does not form a constituent. So if Compositionality holds, either the syntactic analysis in (41) is incorrect, and an alternative must be found for the ellipsis facts, or the semantic analysis of de3 as a nominalizing type-shifter is faulty.

Before we conclude, we consider two other possible solutions. One is that perhaps the examples of putative NP-ellipsis in Chinese in (35)-(37) are not genuine instances of ellipsis. If this were true, the argument from ellipsis for the structure in (41) would bear no weight. The only plausible alternative to ellipsis in (35)-(37) is that these are all free relatives. In fact, Mandarin CP-de relatives, possessives, and attributive measure phrases can occur in contexts parallel to the Thai free relatives in (18):

\[
(42)
\]

\[
\begin{array}{c}
a. \text{ ni qianmian kandao de shi yi dong dalou } \\
\text{ you front see DE COP one CLF building}
\end{array}
\]

‘What you see in front of you is a tower.’

\[
\begin{array}{c}
b. \text{ ni shuo de gen ni zuo de bu yiyang } \\
\text{ you say DE with you do DE not same}
\end{array}
\]

‘What you say is not the same as what you do.’

The availability of ‘free’-de phrases in non-elliptical contexts raises the possibility that (35)-(37) are not elliptical at all. Under such a view, free relatives could occur in elliptical environments and would be able to ‘recover,’ perhaps pragmatically, the semantic content of their antecedent.

An alternative solution is that de in Chinese actually shows its duality in yet another way, namely in what empty categories it licenses. The discussion of ellipsis and free relatives in Chinese and Thai, particularly the fact that Thai allows free relatives (18) but none of the counterparts of ellipsis (35)-(37), suggests that there might be two kinds of empty categories after de: one that occurs after ellipsis (35)-(37), and one that occurs in free relatives (42). The second kind is similar to what occurs in free relative clauses in Thai and is perhaps what Li (2005) calls true empty categories (TEC). Seen through the lens of the Thai data, where the thii operator licenses free relatives and not ellipsis, and assuming that de2 is the equivalent of thii in Thai, we propose the following generalization:
In Chinese, \textit{de2} licenses free relatives/TEC while \textit{de3} licenses ellipsis.

Since \textit{de2} and \textit{de3} are fused into one lexical item in Modern Mandarin, it is expected that their fused form, namely \textit{de}, appears to license a wider variety of empty categories than \textit{thii} in Thai, where the counter part of \textit{de3} is silent and therefore cannot license ellipsis.

Although the last solution is very appealing, we are still not prepared to offer a structural account of \textit{de} that would allow it to perform all its dual functions in semantics while maintaining the integrity of a structure that conforms to the current syntactic theories.

6. Conclusion

The Thai data, along with the dialectal (and historical) data in Chinese, suggest the presence of predicate forming operators in both languages, more robustly in Thai and less so in Chinese. In Mandarin Chinese, this operator, \textit{de2}, is merged with the homophonus \textit{de3} the nominalizer, over time, making the distribution of \textit{de} recalcitrant for a consistent analysis. Coming to terms with this aspect of Mandarin \textit{de} (the fused form equivalent to \textit{de2+de3}) allows us to clarify some of the uncertainties and confusion in the current intense debate about the exact nature of Mandarin \textit{de} and on how to make proper cross-linguistic comparisons between Mandarin and Thai.

There are at least two clear paths for future research. Empirically speaking, the parallel between \textit{de2} and \textit{thii/khɔɔŋ/khanaat} is “constructed” in the sense that the positive data of overt \textit{de2} from dialectal studies all involve adjectival modifiers, not CP/PossP/MeasureP. It would strengthen the link between Mandarin \textit{de2} and Thai \textit{thii/khɔɔŋ/khanaat} if we can find corresponding occurrences of \textit{thii/khɔɔŋ/khanaat} in Chinese dialects either in the form of \textit{de2} or other phonological forms in CP/PossP/MeasureP.

Theoretically speaking, representing Mandarin \textit{de} syntactically has proven difficult. We would like to capture a parallel between semantics and syntax: As Huang and Li (2009) suggest, modification is intersection semantically and coordination syntactically. However, this view must be reconciled with the ellipsis facts in Mandarin Chinese, as conjunction heads typically do not license ellipsis.

References


Liu, Chen-Sheng Luther. 2010. The positive morpheme in Chinese and the adjectival structure. *Lingua* 120.4: 1010-1056.


