Analyticity in the nominal domain: classifiers and the structure of DPs in languages of Southeast Asia

Focus of interest: variation in DP-internal syntax in Southeast Asian languages:
- many language families (Tibeto-Burman, Tai, Austronesian, Austroasiatic), different typological properties
- random and unconstrained ordering of elements?
- often not in line with general headedness/direction of selection:
  e.g. Thai/Khmer head-initial SVO but seem head-final in DPs (due the final occurrence of the heads Num, CL, and Dem); head-final Burmese seems to have a head-initial Dem/D;

(1) Thai, Khmer: N Adj RC Num CL Dem
(2) Burmese: Dem RC N Adj Num CL
(3) Hmong, Malay, Vietnamese: Num CL N Adj RC Dem

- What is responsible for the diversity, and is there a common structural base? How could such a base be identified?


Classifiers as important clues to the internal structure of DPs

(1) The status of classifiers
  - two different views:

  (a) numbers and classifiers instantiate distinct functional head positions, Num and CL (e.g. Pan 1990, Tang 1990):
  (b) numbers and classifiers comprize a single functional head CL/Num/Q
      (e.g. Kawashima 1993, Muromatsu 1998)

Gil (1994), support for position (b):

(i) numbers and classifiers commonly occur together as a single uninterrupted sequence.
(ii) numbers and classifiers in many languages pattern phonologically as a single unit → the classifier is affixed to the numeral
(iii) in some languages, the number+classifier sequence can be separated as a unit from the rest of an NP (Num-CL floating)
(4) nay-ka **chayk-ul** ecey **sey-kwon** ilk-ess-ta
    Korean
    'I read four books yesterday.'

(5) saouq-kou canaw mane-ka thoun-ouq we hke pa te.
    'I bought three books yesterday.'

Support for position (a):

(i) two distinct morphemes → two distinct heads; classifiers in SEA languages are
phonologically unreduced → independent functional words not inflectional affixes

(ii) classifiers are *functional* elements with the primary semantic function of *individuating* NPs
(Muromatsu (1998), Cheng & Sybesma (1999), Li (1998a/b/1999) and others); the
individuation function of classifiers is distinct from number specification provided by
numerals

Effects of (ii): Vietnamese, Hmong, Nung: classifiers occur alone without any numeral, simply
functioning to individuate the NP:

(6) **tus** tssov tshaib tshaib plab
    ‘The tiger is/was very hungry.’ (Hmong, Jaisser 1987)

Numerals occurs without any classifier when numeral specification is vague and individuation is not
(necessarily) implied or focused.

(a) **Nung** and **Burmese**: classifiers optional with numbers which are multiples of ten
(b) **Jingpo** classifiers often omitted with numbers over ten
(c) ‘**Thai**’ classifiers do not occur with large numbers like 1000 unless individuation is implied’
    (Aikhenvald (2000), p.100)
(d) **Malay** classifiers omitted when approximate and vague numeral reference is made, no
    specific individuation Hopper (1986):

(7) adalah **dua tiga pondok** kechil-kechil bersama-sama dekat rumah Temenggong
    be 2 3 hut small small together near house Temenggong
    ‘There were **two or three small huts** close together near Temenggong’s house.

→ CL and Num perform distinct formal functions → separate syntactic heads

Also: **Nung** (Tai): number ‘one’ not adjacent to classifier: CL N Num_{[one]} :

(8) **ahn tahng nuhng** ma
    take **CL chair one** come
    ‘Bring a chair.’ (Saul & Wilson 1964)
Ejagham (Benue-Congo): all numbers non-adjacent to CL:

(9) a-**mege**  \( i \)-cokud  a-**bae**
    NC-CL   GEN   NC-orange seed   NC-two
    \( \text{‘two orange seeds’} \) Watters (1981)

→ distinct functional head positions

Languages where (certain) adjectives occur between Num and CL → Num and CL not a single functional head:

(10) a.  yi  xiao  ben  shu       b.  liu  da  jian  xingli
     one  small  CL  book             6  big  CL  luggage
     \( \text{‘one small book’} \)            \( \text{‘six big pieces of luggage’} \) T’ung & Pollard (1982)

(2) **Headedness and directionality in the DP.**

If Num and CL are separate heads → important consequences for analysis of DP-structure

(11) Thai: \[ DP \text{ N Adj Num CL Dem } \]

(12) Chinese: \[ DP \text{ Dem Num CL Adj N } \]

(13) baan  yai  saam  lang  nii
    house  big  3  CL  this
    \( \text{‘these three big houses’} \)

(14) zhe  san  suo  da  fangzi
    this  3  CL  big  house
    \( \text{‘these three big houses’} \)

If Num and CL a single head \( NC^0 \) → (9) and (10) mirror images of each other → (10) base-generated head-initial DP; (9) base-generated fully head-final DP

(15) a. Thai   DP    b. Chinese   DP
    \[ \begin{array}{c}
    \text{NCP}
    \text{NP}
    \text{house big}
    \end{array} \]
    \[ \begin{array}{c}
    \text{D}^0
    \text{NC}^0
    \text{this}
    \text{3.CL}
    \end{array} \]

    \[ \begin{array}{c}
    \text{DP}
    \text{NCP}
    \text{NP}
    \text{big house}
    \end{array} \]

    \[ \begin{array}{c}
    \text{D}^0
    \text{NC}^0
    \text{this}
    \text{3.CL}
    \end{array} \]

BUT if Num and CL different heads → (11) and (12) cannot both be base-generated

(12) → Num higher than CL; (11) → CL higher than Num
(16) a. 'Thai'?  
\[ \text{DP} \]
\[ \text{CLP} \]
\[ \text{D}^0 \]
\[ \text{this} \]
\[ \text{NumP} \]
\[ \text{CL}^0 \]
\[ \text{cl} \]
\[ \text{NP} \]
\[ \text{house big} \]
\[ 3 \]

b. Chinese  
\[ \text{DP} \]
\[ \text{D} \]
\[ \text{this} \]
\[ \text{Num}^0 \]
\[ \text{CLP} \]
\[ \text{NP} \]
\[ \text{cl} \]
\[ \text{big house} \]

- common assumption: Num higher than CL: nouns first individuated by a classifier, then quantified over by a numeral/other quantifier

→ Thai order in (11) not simply base-generated; in a “pure” head-final DP one expects the ordering: [N CL Num Dem] with Num selecting CLP to its left.
→ must be some movement-distortion in (11)

- 3 possibilities:
  
  (i) the Thai DP is head-final (though Thai is elsewhere head-initial) \[ \rightarrow [N \text{ Num CL}] \] from movement of CL to a higher rightward head position above Num, i.e. \[ [N_t \text{ Num} \ CL_t] \]
  BUT: should violate Head Movement Constraint

(17) Thai head-final?  
\[ \text{DP} \]
\[ \text{XP} \]
\[ \text{D} \]
\[ \text{this} \]
\[ \text{NumP} \]
\[ \text{x}^0 \]
\[ \text{CLP} \]
\[ \text{Num} \]
\[ \text{cl} \]
\[ 3 \]
\[ \text{NP} \]
\[ \text{house big} \]

(ii) \[ [N \text{ Num CL}] \] from lowering of Num to a position below CL, i.e. \[ [N \text{ Num}_t \ CL_t] \]
BUT: barred due to lowering
(18) Thai head-final?

(iii) → assume Num and CL do not change positions → the Thai DP is head-initial and N
moves leftwards from position to the right of CL; as N-raising should be blocked by the
HMC + because adjectives and relative clauses intervene between N and Num\(^0\) CL\(^0\) as in
(1), → assume movement of NP:

(19) a. XP b. [SpecNumP? SpecDP?]

• same conclusion of NP-movement also forced by Dem position: if Thai head-initial + only one
direction of selection per language → DP-final Dem either in D\(^0\) and NP complement raised leftwards,
or DEM in leftward SpecDP and NP raised leftwards:

(20) a. DP b. DP

6
Other head-initial SEA languages with DP-final Dems: **Hmong, Vietnamese, Indonesian.**
→ need DP-internal NP raising with these languages too
PLUS NOTE: these languages are otherwise regularly head-initial inside the DP (Num>CL>NP):

(22)  tiga buah sepeda besar (Indonesian)
     3    CL    bicycle big
     ‘three big bicycles’

(23)      NumP
           Num
             l
    3    CL 0
         l
    cl     NP
        bicycle  big
DP-final Dem is sole exception to head-initial pattern → assume due NP movement

(24)          DP
               NP
       bicycle big   D'  'sepeda besar itu'
             D
       this

• ALSO NOTE: **Nung** (a Tai language related to Thai) has parallel ordering of elements in the DP (and elsewhere) as standard Thai, with the exception of NUM and CL, which **precede** the N:
→ Nung shows suggested underlying order of Thai:

(25)  Nung (northern Tai):  **Num CL** N Adj
     Thai (southern Tai):  N Adj **Num CL**
To allow for both Thai and Nung → need some kind of movement. Most natural to assume NP-movement occurs in standard Thai.

• FURTHER SUPPORT: DP-internal NP-movement occurs in other languages: Indonesian neutral order [Num CL NP Dem] sometimes occurs as [[NP]k Num CL t_k Dem]:

(26) maka adapun mengerjakan [lobang sa buah itu] sampai lima enam hari and indeed make **hole one CL that** took 5 6 day ‘Indeed it took 5 or 6 days just to dig that one hole.’ (Hopper 1986, p317)

Similar alternations in Vietnamese: neutral order [Num CL NP] → [NP_k Num CL t_k ] in poetry/literature and ‘inventory forms’ (Nguyen 1997).

• CONCLUSION: Thai is regularly head-initial in DP. Surface forms result from NP-movement.

**An alternative possibility to NP-movement: NUM-CL as a predicate to NP**

Muromatsu (1998) Num and CL are not functional heads selecting NP, but make up small clause predicates, predicing onto NP subjects within DP.

(27) SC

    NP

                   Q[Num,CL]
                           |
               hon-o          san-satsu
book-ACC        3-CL

→ [NP [Num CL]] structures can be base-generated without movement.

• support for the ‘predicate’ theory: numerals and certain quantifiers can occur as predicates:

(28) a. the apostles were twelve in number
b. they are many (in number)

(29) mie-guo-zhe wu-shi Classical Chinese
destroy-country-NZL 50 ‘His extinctions of countries were fifty.’ (Meng 3B/9)

However, only *weak* not strong quantifiers can occur in subject-predicate structures:

(30) *the men are all/each

But in Thai, Num allows both weak and *strong* quantifiers:

(31) dek soong khon / laai khon / thuk khon
child 2 CL several CL every CL
‘two children/several children/all the children’
as strong quantifiers can occur → cannot be a subject predicate structure
Such strong quantifiers are impossible as predicates in Thai:

(32) *dek-law-nii thuk-(khon)
    child-group-this every-CL
    intended: ‘*These children are all/every.’

→ assume DP-internal NP-movement analysis instead

(3) What causes DP-internal NP-movement?

• Greenberg (1975), Gil (1994) and others: sequences of N/NP > numeral/classifiers often occur in written list or ‘inventory’ forms + when people are involved in situations such as buying commodities in a store and placing orders for food in a restaurant.

Generalization: if both [Num CL NP] and [NP Num CL] are possible in a language (e.g. in Indonesian, Vietnamese) → [NP Num CL] occurs preferred in lists or buying/food-ordering situations and often only occurs in such situations.

NB: [NP Num CL] also occurs as a conventional way of itemizing elements in written shopping lists in English which otherwise do not permit such orders, as shown in (34):

(33) saya mau membeli beras dua kilo
    I want buy rice 2 kilo
    ‘I want to buy two kilos of rice.’

(34) Sugar, 3 pounds
    Bread, 2 loaves
    Wine, 4 bottles

[NP Num CL] orders are very widespread, and occur throughout the major classifier languages in Asia, either as the ONLY possible ordering (Thai, Khmer, Burmese), or as alternatives to [Num CL NP] orders in certain situations (Indonesian, Vietnamese, Chinese, Japanese, Korean).

Why?

• Possible functional answer: placing NP before Num/CL is a natural/useful way of ordering this information in certain types of presentational situations → information about the identity of the NP is presented before information about its cardinality.

Such ordering is practical/useful at certain times, e.g. a storekeeper receiving information about what goods a customer wishes to purchase: identifying the type of goods before the quantity (i.e. ordering NP before NUM/CL) presents the information in a sequence which mirrors the actions of the storekeeper, who first needs to identify and locate the required goods and then select a certain quantity of them. Presentation of the information in this way is helpful and efficient/logical.
placement of the NP in DP-initial position is like **presentational focus** or **topicalisation**

- in both CP and DP, newly presented nominal elements may be fronted so that they precede additional information added on about them.

**Classical Chinese**: had both [Num NP] and [NP Num CL] forms. [NP NUM CL] occurs in initial presentation of the new referent in (23), [Num NP] without NP-fronting follows when the referent has been established as old, identified information:

(35) you da [jiang] er ren. Er [jiang]…
    be big [general] 2 CL 2 [general]..
    ‘There were two great generals. The two generals…’ (Schafer (1948, p.413)

- Thai, Khmer and Burmese no alternative to [NP Num CL] order → suggest that “presentational focus” has become grammaticalized as obligatory (similar to the grammaticalization of XP-fronting in German main clauses)

**NEXT QUESTION**: So, did modern [NP Num CL] develop from earlier [[Num CL NP] through NP-fronting?

**ANSWER**: No, not as far as we know.

**FOLLOW-ON QUESTION**: So how can we explain the “sudden development” of [NP Num CL]?

A possible explanation lies in considering a different theory of [NP Num CL], the ‘adverb’ theory.

**The adverb theory of NP Num CL**

Num CL are *not* inside the DP but are adverbs base-generated in distinct adjunct-like positions. Fukushima (1991), Ishii (2000) Greenberg (1975). Due:

(a) Num CL can occur separated from NP in other VP-/S-final adverbial positions:

(36) mii [nisit] maa haa [khun] [soong] [khon] Thai
    be student come find you two [CL]
    ‘Two students came to look for you.’

(37) xi [sang] [di] yu Qin [qi-bai-li] Chinese
    west lose land to Qin 700 [li]
    ‘On the west we lost 700 li’s land to the Qin’ (Meng 1A/5, in Pulleyblank 1995)

(b) Before the formal establishment of the functional category “Classifier/CL”, “classifiers” would have been the heads of independent NPs occurring as adverbial type elements modified by numerals:

(38) [NP ren] [NP shi-you-wu ren] person 10-and-5 person
    ‘15 people”
(c) NUM and CL in Japanese can be co-ordinated with adverbs:

(39) Shoonin-ga kinoo [san-nin katsu tashika-ni] sono jiko-o mokugekishi-ta
    witness-NOM yesterday 3-CL and certainly that accident-ACC witness-PAST
    ‘Three witnesses certainly witnessed that accident yesterday.’ (Fukushima 1991)

(d) Instances of ‘floated’ NUM-CL pairs which could not have been base-generated with the
    associated preceding NP → must be base-generated independently:

(40) a. Taroo-ga pisutoru-o ippatsu kinoo utta
    Taroo-NOM pistol-ACC one-CL(blast) yesterday shot
    ‘Yesterday Taroo shot off one blast of his pistol.’ (Fukushima 1991)

b. *Taroo-ga ippatsu-no-pisutoru kinoo utta
    Taroo-NOM one-CL-GEN-pistol-ACC yesterday shot

• **Proposal**: early sequences of NP and numerically-quantified adverbial NP were re-interpreted as
  parts of a single DP. When this happens, the NP>Num CL ordering results in an instantaneous
  analysis of movement, as this is the only analysis consistent with speaker’s head-initial parameter of
  selection. The natural linear ordering of NP > Num CL present in NP + adverb structures is re-
  interpreted as DP-internal presentational focus movement. Movement arises instantaneously from
  structural re-analysis.

(41) early: [NP NP] [NP Num N] is reanalysed as: [[NP NP] Num CL NP ]

**NEXT PROBLEM**: It now needs to be proved that re-analysis of NP + adverb Num-CL really has taken
place and that NP Num-CL is not still NP followed by ADV, as argued in Fukushima and Greenberg.

**EVIDENCE**: (a) in Thai & Khmer, Dem follows Num and CL → Num and CL are DP-internal:

(42) [DP dek saam khon nau] keng maak
    child 3 CL Dem clever very
    ‘Those three children are very smart.’

(b) Burmese case-markers & postpositions follow Num & CL → Num & CL are DP-internal, not
    adverbial:

(43) canaw [DPsaouq hna ouq]-ko weh hta teh
    I book 2 CL ACC buy ASP REALIS
    ‘I bought two books.’

(44) U-Win-Win [FP [DP meitswee thoun yauq] ne] Yangoun thwaa teh
    U-Win-Win friend 3 CL with Rangoon go REALIS
    ‘U-Win-Win went to Rangoon with three friends.’
Aspects of the interpretation of [NP Num CL] sequences → Num & CL are DP-internal. Example (45) Num and CL in S-final adverbial position → partitive interpretation:

(45) **dek** sia chiwit laew **saam khon**
Thai
child lose life ASP 3 CL
‘Three of the children died already.’

(46) NP NUM CL sequence → *no* partitive interpretation possible

(46) **dek** **saam khon** sia chiwit laew
Thai
child 3 CL lose life ASP
‘The three children died already.’

Partitive readings assumed possible when numerical quantification is applied to a definite DP from a DP-external position (47a); partitive readings are blocked when numerals occur under the scope of D₀ inside the DP, (47b):

(47) a. [QP Three of [DP the children]].
   b. [DP The three children]

Thai (45): Num & CL DP-external → partitive reading
Thai (46): Num & CL adjacent to NP → no partitive reading possible → conclude Num & CL in [NP Num CL] are DP-internal.

Similar evidence in Burmese: (48) [Dem NP Num CL] only a non-partitive interpretation → Num and CL are DP-internal under the scope of the D₀:

(48) canaw eh-dii saouq hna ouq weh hta teh
I Dem book 2 CL buy ASP REALIS
‘I bought those two books/*I bought two of those books.’

Cross-linguistically, not all [NP Num CL] forms pattern in the same way. [Dem NP Num CL] in Japanese (49) similar to Burmese (48) may have the opposite partitive interpretation to Burmese:

(49) Jiro-wa sono hon-o san satsu katta (Muromatsu 1998)
   Jiro-TOP DEM book ACC 3 CL bought
   ‘Jiro bought three of those books./NOT:*Jiro bought those three books.’

→ Japanese Num CL may not be inside DP and might be adverbial as argued in Fukushima (1991) and Ishii (2000). Burmese may also still allow for such an adverbial possibility. Case-markers often follow Num CL in Burmese, but can also occur between NP and Num CL and (50) → same partitive-like interpretation as Japanese (49) → still possible that Num CL in Burmese occurs outside of the DP:

(50) U-Win-Win eh-dii saouq-ko hna ouq weh teh
    U-Win-Win Dem book ACC 2 CL buy REALIS
    ‘U-Win-Win bought two of those books.’
The historical development of Khmer [NP NUM CL]

(51)     Old Khmer       Middle Khmer       Modern Khmer
a.      N Num
b.      N CL Num       N CL Num

c.      N Num CL       N Num CL

Possible explanation of the development:

(a) old Khmer: early ‘classifiers’ were really just nouns →
N CL Num = sequences of two NPs: [NP N₁] [NP N₂ Num] (noting that [NP N Num] exists independently in Old Khmer - (51a))

(b) middle and modern Khmer: grammaticalisation of N₂ as the new category classifier → two-NP structure is re-analysed as one NP/DP, Num and ‘CL’ internal to DP causes important change in CL/Num word order, reversing linear order of CL and Num pattern (c) replaces pattern (b):

(52)     [NP N₁] [NP N₂ Num] → [N CL Num] triggering change to [N Num CL]

→ when grammaticalisation and re-analysis occurred → forced Num and CL as functional heads within DP to be re-aligned in a head-initial order following the general direction of headedness in Khmer

• if correct → shows strong pressure for languages to adopt consistent head-initial orders
→ grammaticalisation and reanalysis not random; there are clear principles of headedness governing the organisation of DPs.

CONCLUSIONS:
• DP-internal NP-movement as a significant cause of word order variation
• NP-movement is the result of reanalysis of an earlier adverbial form
• adverbial Num-CL forms may still exist in certain languages
• the switch in CL/Num order in middle Khmer is understandable once one adopts such an account of the development of classifiers in DPs

SECTION 2: DP-internal X₀-movement and grammaticalization

How X₀-movement and grammaticalisation may be responsible for further surface variation in DPs in SEA languages.

(4) Bare classifiers and definiteness

• an interesting pattern found in various SEA languages – bare classifier-NP sequences with no Num:
• an important general observation: bare classifier-NP/CL-NP sequences are commonly associated with definiteness effects; when a DP has a definite interpretation, a bare classifier is generally found to occur with it in these languages, similar to Cantonese

• Cheng & Sybesma (1999): classifiers in Cantonese are definite elements like determiners in other languages; classifiers however occur in CL so → CL-NP with definite interpretations are simply CLPs and do not project any higher functional structure such as NumP or DP

• a different approach: bare classifier structures are regular DPs in which CL raises up to D

• to be accounted for: although CL-NP may be definite, when Num occurs → can only be indefinite:

(56) toi mua tam cai ghe (Vietnamese)
I buy 8 CL chair
‘I bought eight chairs.’

→ unexpected if classifiers are inherently definite; expect (56) = either ‘the eight chairs’ or ‘eight of the chairs’

→ need to capture the fact that classifiers are definite at some times but not at others

• also: when indefinite [Num CL NP] is combined with a demonstrative → definite interpretation:

(57) goh saam bo sue (Cantonese)
Dem 3 CL book
‘those three books’

(58) slong ahn sleng te (Nung)
2 CL province Dem
‘those two provinces’

→ the locus of definiteness is the D position above NumP + the occurrence of a particular morpheme in either D or SpecDP → interpretation of the DP as definite
**Suggestion:** bare classifier [CL NP] forms are definite due to (optional) CL raising up to D\(^0\) either D\(^0\) or SpecDP must be overtly instantiated for the definite interpretation to be triggered otherwise the DP is interpreted as indefinite by default

- when Num occurs, no possible CL-to-D due to HMC, hence no: [*CL\(_i\) Num \(_t\) NP] instead Dem has to be inserted into D\(^0\) (or SpecDP) to trigger a definite interpretation

\[(59)\]

\[
\text{DP} \quad \begin{array}{c}
\text{D} \\
\text{NumP} \\
\text{Num} \\
\text{CL} \\
\text{NP}
\end{array}
\]

- potential empirical evidence for the higher D position: **two CL structures** in certain varieties of Vietnamese, a second general classifier element may occur before the regular classifier and results in sequences with clear definite interpretations (Nguyen 1997):

\[(60)\]  

\[\text{cai con dao [anh cho toi muon], no that sac} \]
\[
\text{CL CL knife you give me borrow, it real sharp} \\
\text{‘The knife you gave me is really sharp.’}
\]

\[(61)\]  

\[\text{cai chiec ban nay} \]
\[
\text{CL CL table Dem} \\
\text{‘this table’}
\]

→ shows that there is another X\(^0\) head-position above CL\(^0\) + this X\(^0\) is (a) instantiated by a classifier and (b) specifically associated with definite interpretations of the DP

- **grammaticalization of CL\(_{\text{general}}\) in D\(^0\)**

  suggestion: the possibility of inserting CL\(_{\text{general}}\) directly into D\(^0\) is due to movement and reanalysis. Sustained CL-to-D movement → CL\(_{\text{general}}\) allows for potential reanalysis as a D\(^0\)-element → allows CL\(^0\) to be lexicalised and instantiated by a second classifier as in (60-61)


  → synchronically, if CL\(_{\text{general}}\) and a second regular CL are selected in the numeration → CL\(_{\text{gen}}\) is inserted into D\(^0\), and no CL-to-D movement (45-46)

  if CL\(_{\text{general}}\) not selected together with a second regular CL → the classifier base-generated in CL\(^0\) undergoes CL-to-D to trigger the definite interpretation
(5) Num-raising and indefiniteness

• further surface variation from head-movement: the patterning of number ‘one’ in Thai and Nung
• closely-related Tai languages; the regular position for Num is before CL (62)
• however, number one is commonly placed following CL in Thai and only occurs DP-finally in Nung (63a/b):

(62)  a.  dek  saam khon  (Thai)       b.  slam ahn  vet  (Nung)
      child  3  CL               3  CL  spoon
     ‘three children’            ‘three spoons’

(63)  a.  dek  khon nung  (Thai)       b.  ahn  tahng  nuhng  (Nung)
      child  CL one               CL  chair  one
     ‘one/a child’                ‘one/a chair’ (Saul & Wilson 1980)

• both Thai & Nung (re-)positioning of number ‘one’ → ‘one’ now commonly placed in the DP-final position that demonstratives/DEM occur in (64):

(64)  a.  dek song khon nii  (Thai)       b.  slong ohng dehk  te  (Nung)
      child 2  CL  Dem             2  CL  child  Dem
     ‘these two children’          ‘those two children’

Thai (65a) → ‘one’ in DP-final position is in complementary distribution with Dem
‘one’ can otherwise occur in the pre-CL Num0 position with a Dem following CL (65b):

(65)  a.  *dek  khon nii  (Thai)       b.  dek  nung khon nii
      child CL  one  Dem         child one  CL  Dem
     ‘this one child’             ‘this one child’

→ ‘one’ in Thai & Nung is coming to be an indefinite determiner which contrasts in its indefinite specification with the definiteness encoded by demonstratives

The Syntax of Num-to-D Raising

• sections 2/3 argued for NP-movement over Num & CL in Thai, i.e. [NP_i Num CL t_i ]
• section 2 also argued that elements to the right of Dem in Vietnamese, Nung, Indonesian, Thai & Khmer raised to the left of Dem as a single constituent: [[Num CL NP], Dem t_i ]
→ in Thai & Khmer when a Dem occurs → there will be two applications of leftwards XP movement:

1 If no NP over Num CL raising occurs, the surface order is that of Vietnamese, Nung, and Indonesian, and if NP-movement but no movement over Dem occurs, the order is that found in Burmese.
underlying structure:  [ Dem Num CL NP ]

NP movement:  [ Dem Num CL t ]

movement over Dem:  [[NP Num CL t kl Dem tk] ]

What are the landing-sites for the two movements? - Three possibilities.

(i) NP moves to SpecNumP, [NumP NP Num CL] moves to SpecDP

(ii) NP left-joins to NumP, [NumP NP Num CL] left-joins to DP

(iii) Movement of NP and [NumP NP Num CL] are both instances of DP-internal topicalisation, to the specifier positions of DP-internal topic projections below and above D0; similar to the assumption (Rizzi (1997)) that there may be iterated topic positions in the higher part of clausal structure.

Possibility (i):

(67)

Possibility (ii):

(68)
'one' raising

- ‘one’ is base-generated in Num\(^0\) and raises up to D\(^0\)
  - note that it is not possible to have any other number in Num when ‘one’ is in the DEM position:

(69)  *dek saam/nung khon nung
      child 3/one  CL  ‘one’

THE DERIVATION:

(70)  underlying structure:
      [ one  CL  NP ]

NP movement over Num CL:
      [ NP\(_k\) one CL  \( \tilde{t}_k \) ]

‘one’ raises from Num\(^0\) to D\(^0\):
      [ one\(_i\) [ NP\(_k\) \( \tilde{t}_i \) CL  \( t_k \) ] ]

NumP raises over D\(^0\):
      [ NP\(_k\) \( \tilde{t}_i \) CL  \( t_k \) ]\(_m\) [ one\(_i\) \( \tilde{t}_m \) ]
• three conclusions

• good support for the conclusion that the DP in Thai is not head-final; if the DP was head-final and ‘one’ raised rightwards to $D^0$ over CL, this should violate the HMC (74); Num-to-D raising can only be possible if Num is higher than CL → the Thai DP is head-initial → NP to the left of Num CL has raised there from a position to the right of CL

(74) 'head-final Thai'

• further support for the assumption that $Num^0$ and $CL^0$ are independent functional heads - if ‘one’ raises out of the regular position of numerals and up to $D^0$ → classifiers are not suffixes attached to numerals in a single head; affixes/subparts of words are not normally stranded by operations of movement; instead → Num and CL are independent words in discrete functional heads; see tree (75);
Illicit extraction of a sub-part of a head if Num + CL form a single head 'NCP':

\[(75)\]
\[
\begin{array}{c}
\text{DP} \\
\text{\textbf{\small D}} \\
\text{\textbf{\small one}} \\
\text{\textbf{\small NP}} \\
\text{\textbf{\small house big NC}} \\
\text{\textbf{\small NC'}} \\
\text{\textbf{\small NP}} \\
\text{\textbf{\small one-CL}}
\end{array}
\]

- the Thai and Nung new indefinite determiner 'one' targets the same position as definite D\(^0\) elements (demonstratives) \(\rightarrow\) supports the assumption that both definiteness and indefiniteness are encoded in the same position D\(^0\) in DPs \(\rightarrow\) indefinite nominal expressions are full DPs in the Tai languages.

Aside: the number 'one' may also raise to D\(^0\) in English. In contrast to other numerals, 'one' cannot co-occur with other instantiations of D\(^0\) unless it is interpreted like the adjective 'unique' and is followed by a relative clause as in (76):

\[(76)\]
\[
a. \text{The two/three men/*one man left.} \\
b. \text{These two/three books are mine/*this one book is mine.} \\
c. \text{The unique/one person *(I like) is David.}
\]

(6) Head-movement below D\(^0\) and Num\(^0\): N-to-CL

- further classifier-related variation - 'classifier-less' DPs
- no classifier occurs with a numerically-quantified NP although individuation is implied
- two basic forms:
  - (a) in some language certain nouns are commonly quantified without any CL;
    frequently e.g. 'year', 'day', 'time' etc, sometimes 'person', + other high frequency terms:

\[(77)\]
\[
a. \text{hai nam} \quad \text{(Vietnamese)} \quad b. \text{ba lan} \\
2 \text{ year} \quad \text{3 time} \\
'\text{two years}' \quad '\text{three times'}
\]

  - (b) in certain languages (e.g. colloquial Minangkabau and Indonesian) the use of a CL is optional with a fully wide range of nouns, e.g. in Nung:

\[(78)\]
\[
\text{slam (tew) kha-lo} \quad \text{(Nung)} \\
3 \quad \text{(CL) road} \\
'\text{three roads'}
\]
• three possible analyses

I. No CL position in type (a)? → some nouns require no individuation?
Problem: will not generalize to cover type (b) where the nouns do have classifiers, these are just optional.

II. Type (a) elements ‘year’ and ‘time’ are base-generated in CL and there is no N/NP?
Unlikely – if no N-position → there are classifiers which classify/individuate nothing, but classifiers are by definition functions which apply to a second complement term → the presence of a CL requires the presence of an N/NP; also → will again not generalise to cover type (b).

III. If there is Num-to-D and also CL-to-D, possibly there is also N-to-CL:
the elements ‘year/time’ are base-generated in N⁰ and raise to CL instantiating both positions

(79)
\[
\begin{array}{c}
\text{DP} \\
\text{D} \\
\text{NumP} \\
\text{Num}^0 \\
\text{CLP} \\
3 \\
\text{CL} \\
\text{NP} \\
\text{N} \\
\text{year}
\end{array}
\]

• CL and N are linearly adjacent in Vietnamese, Indonesian, Chinese → difficult to see N-to-CL
• Thai & Khmer: different surface linearisation of DP structure → N and CL are not adjacent but separated by numerals in Num due to NP movement: (i.e. [N(P) Num CL]); → possible to check the N-to-CL hypothesis;

(80)
\[
\begin{array}{c}
\text{DP} \\
\text{D} \\
\text{NumP} \\
\text{NP} \\
\text{house big} \\
\text{Num}^0 \\
\text{CLP} \\
\text{one} \\
\text{CL}^0 \\
\text{NP} \\
\text{house big}
\end{array}
\]
(81) shows that elements such as ‘year’ and ‘time’ appear in CL to the right of Num, so not in \( N^0 \) → indicates N-to-CL movement:

(81)  a. soong pii  b. saam khrang  c. sii khon (Thai)
     2 year                   3 time                   4 person
     ‘two years’               ‘three years’             ‘four people’

Derivation of N-to-CL in Thai:

(82)  a. NumP  b. NumP
    Num
    |  CLP
    two  CL  NP
    |   year  N
    |   |  year

• the N-to-CL analysis suggests that certain nouns may be able to function as their own classifiers and lexicalize CL\(^0\) as well as N\(^0\). This possibility is seen in Thai, Burmese, Lahu and other languages in a second way with the occurrence of “self-classifiers” or “repeaters” – the repetition in CL of the element in N:

(83)  a. hoong saam hoong (Thai)  b. cun ta cun (Burmese)
     room 3 room/CL           island one island/CL
     ‘three rooms’             ‘one island’

→ the idea of an element being used to classify/individuate itself is plausible and commonly attested.

• if one assumes the ‘copy theory’ of movement (Chomsky 1995), → movement between N and CL and the copy left in N is not deleted/is spelt-out phonetically in repeater cases (though the copy of NP-movement would be deleted, otherwise giving rise to a redundant, identical sequence of elements):

(84)  a. NumP  b. NumP
    Num
    |  CLP
    three  CL  NP
    |   room  N
    |   |   room
    |   |   |   room
    |   |   |   |   room
• the N-to-CL raising analysis of type (a) will extend to cover type (b) if a CL is not selected from the lexicon there will be N-to-CL; if a CL is selected, there will be no N-to-CL (85) represents Nung (78);²

(76) a. overt CL not selected  b. overt CL selected

![Diagram](image)

• finally, if N-to-CL occurs regularly, and if the ‘movement and reanalysis’ approach to grammaticalisation is correct, one might expect certain ‘N’’s might become reanalyzed as CLs; cross-linguistically CLs are indeed known to develop from Ns;

(7) Conclusions

• surface variation in DP structure is the result of a combination of XP and X₀ movement inside the DP
• X₀ heads may commonly raise to instantiate/lexicalize more than one position
• regularization of (head-)raising leads to grammaticalization and categorial reanalysis

References

---

² In this sense classifiers pattern a little like expletives – when an expletive such as ‘there’ is selected in a numeration, it will be inserted into SpecIP, but when no such expletive is selected, the subject of a clause has to raise to instantiate the SpecIP position.


Pan, Hai-Hua. (1990) Head noun movements in Chinese, the ECP and minimality. Ms. U. of Texas at Austin.


