Classifiers, Demonstratives and Classifier-to-Demonstrative Movement[*]

BRIAN, HOK-SHING, CHAN

Abstract

This paper examines the feature composition of Cantonese determiner-like elements in the light of current literature (Cheng 1998, Cheng and Sybesma, to appear, Li 1998, 1999a, b). It is argued that a combination of proposals in these works best captures the Cantonese facts. In particular, I argue for three specific positions: Firstly, classifiers, rather than demonstratives, are instantiations of D with a [referential] feature. Secondly, classifiers encode a [count] feature. Third, demonstratives encode a [definite] feature, but classifiers and numerals do not. What appears to be a bare classifier phrase in Cantonese (i.e. without an overt numeral or demonstrative) is actually bound by a Dem node, which may be empty (resulting in an indefinite reading) or filled by the classifier which moves from the lower CL position.

1 Introduction

Following Abney (1987), it has been standardly assumed that nouns not only project *noun phrases (NP's)* but also *determiner phrases (DP's)*, of which the *determiner (D)* is the (functional) head. More recently, D has beem considered an operator position which binds an NP and turns it into an argument (Stowell 1989, 1991, Longobardi 1994, 1996). Taken seriously, this assumption amounts to the claim that D is obligatorily present in the "superstructure" of nominal expressions across all languages. This idea of D as a universal construct is consistent with a restrictive theory of phrase structure in which the underlying phrase structure for all languages is identical (Kayne 1994, Zwart 1997). On the other hand, whereas Abney (1987) appears to consider various elements to be in D (e.g. articles, demonstratives, numerals, quantifiers), recent literature suggests that only the *articles* (e.g. "*a/an/the*" in English) occupy the D position (Giusti 1997, Ritter 1995, Szabolcsi 1994).

[*]I would like to thank Professor Neil Smith, Dr Ad Neeleman and Professor Richard Hudson for their insightful comments on previous drafts. All errors are mine.

This version of the DP hypothesis runs into problems in the face of many article-less languages in the world. This paper considers one of these languages, namely, Cantonese. There are a number of possibilities as regards the superstructure of NP's in Cantonese. Let's consider two proposals first.

Covert-D Hypothesis: D might be projected and remain covert alongside other determiner-like elements in Cantonese (i.e. demonstratives, classifiers and numerals). Progovac (1998) presents some evidence which suggests the presence of an empty D slot in Serbo-Croatian, a language without overt articles. The following is some of the relevant evidence:

(1) a. I **samu Mariju** to nervira/*I **Mariju samu** to nervira
and alone Mary that irritates
'That irritates even Mary'

b. I **nyu/mene samu** to nervira/*I **samu nyu/mene** to nervira
and her/me alone that irritates
'That irritates even her/me'

(Serbo-Croatian, Progovac 1998: 167, (7)-(10))

In the above examples, a proper name, which Progovac (1998) controversially considers a *noun*, must follow the adjective in (1a), but a pronoun must precede the adjective in (1b). To account for the word order of these patterns, Progovac concludes that the proper name in (1a) remains an N *in-situ*, whereas the pronoun in (1b) surfaces in D. In (1a), the noun remains *in-situ* because the [referential] feature is weak in D for Serbo-Croatian as in Germanic languages, as suggested by Longobardi (1994). The noun raises to D only covertly because of economy, namely, the principle of *procrastinate*. As for (1b), the pronoun surfaces in D presumably because it is generated there.[1]

There is no similar evidence that Cantonese projects a covert D node higher than N. In the following Cantonese examples (2), we can see that a pronoun follows its modifier[2] [3,]

[1] Progovac (1998: 170, see also her fn.3) actually assumes that pronouns are base-generated lower than D and raised to D eventually. Nonetheless, it is unclear what motivates the (overt) movement of pronouns, given that the [referential] feature in D is assumed to be weak.

[2] Some scholars (e.g. Tse 1990: 75) have observed that pronouns are modified in written Mandarin Chinese. They suggest that modification of pronouns is due to "Europeanization" of Chinese (primarily due to English). Even granted that this is the case, the fact that the modifier is prenominal is distinctively Chinese: English pronouns are rarely modified, and in those rare cases the modifier is usually postnominal:

and does not surface in a higher position as in the case of Serbo-Croatian. Indeed, Cantonese nouns and pronouns are strictly *head-final* in NP's - There is simply no evidence of overt N movement, which has been considered a clear proof that there are functional heads/projections above N (Cinque 1994, Longobardi 1994, Progovac 1998).

(2) **sap6 nin4 cin4 ge3 ngo5** ji5 ging1 sei2 zo2 laa3[4]
ten year before P me already die ASP PRT
'I am no longer what I was ten years ago'
(Literally: 'The "me" ten years ago died already.')

There is only one remaining way to preserve the position that Cantonese has a D node - by claiming that pronouns moves to D covertly. In that case, strong independent evidence is called for since the movement is not visible. Such evidence seems to be lacking in Cantonese.

NP hypothesis: Ernst (1991), Gao (1994) and Lin (1997) argue that there is no D in Chinese nominal expressions. Nouns only project to *noun phrases*, with the functional elements forming a *classifier phrase*. That is, the classifier heads a classifier phrase which is in turn the specifier of the head noun.

---

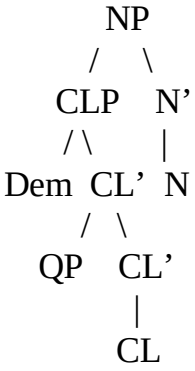
i. He who hesitates is lost. (Quirk et al. 1985: 352)

[3] Cheng and Sybesma (to appear: 28, (60)) give a similar example in Mandarin, in which a pronoun is pre-modified.

[4] The transcriptions of Cantonese follows "the LSHK Cantonese Romanization Scheme (LSHK - Linguistic Society of Hong Kong)", whereas those of Mandarin Chinese follows "Hanyu Pinyin (Chinese Romanization)" which is the standard romanization system in China.

Notations: ASP - aspect marker, CL - classifier, DEM - demonstrative, EMP - emphatic marker, PASS - passive marker, NUM - numeral, NEG - negation, P - preposition, PL - plural marker, PROG - progressive, PRT- particle, Q - interrogative marker.

(3)



(Lin 1997: 419)

This "CLP-as-specifier" analysis in (3) deviates from the standard assumption that the extended projection (in the sense of Grimshaw 1991) of a lexical head is a functional projection (DP, IP). In such a structure, the classifier phrase is a specifier and an optional element. It follows that in those cases where there are no overt classifiers the classifier node need not be projected (This is in line with Chomsky's (1995b) *Bare Phrase Structure Theory*). A non-trivial problem then arises as to where a structure without a functional head (i.e. NP) carries the "functional" information which contributes directly to semantic interpretation.[5] Recent literature suggests that a D is an obligatory operator position which contributes to the referential interpretation (i.e. definite/indefinite/generic) of an argument (Longobardi 1994, Stowell 1989). Under Chomsky's (1995a) minimalist theory, such information, though not lexically realized, has to be present in the syntactic structure so that the meaning of the sentence can be read off from syntactic structure at LF.

Syntactic evidence for a functional head position above N can be drawn from the constraint on the distribution of indefinite "bare nouns" in Cantonese:

(4) ngo5 zong6 dou2 **pang4 jau5**
I meet ASP friend
'I've met a friend.' ("*pang4 jau5*(friend)" is indefinite)

[5] It seems to me that the above-mentioned authors who argue for the "CLP-as-specifier" analysis have not addressed this problem.

(5) **naam4 jan2** hou2 jiu3 min6 zi2
man EMP want face
'Men very much want to have face.'/
*'There are some men who want to have face.'
("*naam4 jan2*(men)" is generic)

In the spirit of Longobardi (1994) [6], Cheng and Sybesma (to appear) assume that the indefinite interpretation is yielded when a bare noun is dominated by an empty CL position where classifiers are base-generated. The asymmetry in (4) and (5) is explained as follows: In object position, there is an empty CL position dominating the bare noun, contributing to the indefinite interpretation. Alternatively, a noun may move to CL, yielding a generic interpretation (e.g. (5)).[7] On the other hand, in subject position, the bare noun has to move to CL: If the bare noun *remained* in-situ and *left* the higher CL position empty, the empty CL node would not be lexically governed and the resultant sentence would violate the *Empty Category Principle/ECP*. Note that the ECP does not drive movement - It is an independent principle which rules out ungrammatical structures, or "crashing derivations" in Chomsky's (1995a) terms.

It is unclear how the "CLP-as-specifier" analysis in (3) can capture the asymmetry in (4) and (5): The bare nouns in (4) and (5) would be simply NP's if the "CLP specifier" is not projected. The constraint on indefinite "bare nouns" in subject position would have to be accounted for by independent principles (e.g. *the Definiteness Constraint* - L-J. Xu 1995). Even granted that a "CLP specifier" is obligatorily projected, while the object "bare nouns" are licensed by the ECP, the ECP would rule out any NP from subject position, which is obviously false. Notice that the bare noun in subject position cannot

[6] Longobardi (1994: 616, (14a-c)) noted a similar asymmetry in Italian.

[7] Note that the movement of N-to-CL in (5) does not led to a "noun-adjective"order, as such order is not possible in Cantonese. For instance,

i. kan4 lik6 ge3 hok6 sang1 tai2 hou2 do1 caam1 haau2 syu1
diligent LNK student read EMP many reference book
"Diligent students read a lot of reference books."

Cheng and Sybesma (to appear: 12) conclude that N-to-CL movement is covert. Nonetheless, notice that an adjective in Cantonese canonically appears in the specifier position of "*ge3*", a functional head (Cheng 1998), contrary to other languages like English. The reason why Cantonese adjective canonically appear in the specifier position of a functional head is still unclear. In any case, the canonical "adjective-noun" order in Cantonese (as in (i)) does not necessarily indicate that the proposed N-to-CL movement in (5) must be covert.

move to Spec-NP in subject position - as we do not normally assume the possibility of head-to-specifier movement.

2 Recent works (Cheng 1998, Cheng & Sybesma, to appear, Li 1998, 1999a, b)

We have just seen that neither the Covert-D-Hypothesis nor the NP Hypothesis is desirable. The remaining possibility is that D is filled by some determiner-like element in Cantonese, despite the fact that this element is not an article. Two obvious candidates are the demonstrative and the classifiers.[8] It is interesting to note that recent works on the superstructure of Chinese nominal expressions differ on this issue (Cheng 1998, Cheng and Sybesma, to appear, Li 1998, 1999a, b)[9]. Li (1998, 1999 a, b) considers demonstratives to be in D, while Cheng and Sybesma (to appear) treats classifiers as D. On top of that, these authors also differ from each other on other topics: Li (op. cit.) considers that demonstratives carry a [+definite] feature, whereas Cheng (1998) and Cheng and Sybesma (to appear) treat a CLP as inherently definite and a NumP as inherently indefinite. Furthermore, Li (1999a, b) considers numerals to be in Spec-NumP and Num is specified for a [plural] or [singular] feature. On the other hand, Cheng and Sybesma (to appear) suggest that classifiers express grammatical number in Chinese, but they do not make a concrete proposal of how that works.

In the following sections, I attempt to show that a combination of the proposals by Cheng and Sybesma (op. cit.) and Li (op. cit.) works best to capture the Cantonese facts. I will argue for the following three positions:

- Classifiers, rather than demonstratives, are instantiations of D (with Cheng and Sybesma, to appear, contra Li 1998, 1999a, b)
- Grammatical number is expressed by classifiers which carry a [count] feature (elaborating Cheng and Sybesma, to appear, contra Li 1998a, 1999a, b)
- Demonstratives carry a [+definite] feature (Li 1998, 1999a, b). Neither classifiers nor numerals are specified for a [definite] feature (contra Cheng and Sybesma, to appear).

---

[8] Numerals are standardly considered to be in the specifier of NumP (Giusti 1997, Haegeman and Guéron 1999).

[9] Cheng (1998) and Cheng and Sybesma (to appear) focus on Cantonese and Mandarin Chinese, whereas Li (1998, 1999a, b) concentrates on Mandarin Chinese.

3 Classifiers, not demonstratives, are instantiations of D

Intuitively speaking, demonstratives are a good candidate to fill D, since they are [+definite], just like the definite article which is assumed to be in D.[10] J. Tang (1990) and A. Li (1998, 1999a, b) suggest that Chinese demonstratives ("*zhe*(this)" and "*na*(that)") are instantiations of D. Assuming that D itself is sufficient to bind an NP (Giusti 1995, Osawa 1998), there arises one immediate objection to this analysis: In Cantonese, the demonstrative alone is not sufficient to bind NP, as shown in the following:

(6) go2 *(go3) jan4
DEM CL person
'that person'

In a cross-linguistic perspective, it is also unlikely for demonstratives and definite articles (which are assumed to be in D) to be base-generated in the same structural position (Bernstein 1997, Burgè 1996, Ernst 1991, Giusti 1997, Szabolcsi 1994, and the references cited there), as demonstratives co-occur with definite articles in many languages, including Modern Greek, Hebrew, Irish, Chamorro, Romanian, Hungarian and Javanese. For instance,

(7) afto to oreo to vivlio
this the good the book
'this good book'

(Modern Greek, Giusti 1997: 109, (41a))

There are other differences between demonstratives and the definite article, as pointed out by Bernstein (1997: 93): In many Romance and Germanic languages a demonstrative can stand alone as a pronoun but the definite article cannot (e.g. English).

---

[10] Bernstein (1997) points out that demonstratives may be indefinite but specific in spoken English. The following is her example:

i. There's this book (that) you ought to read. (Bernstein 1997: 95, (20))

Such an indefinite, specific reading does not seem to exist in Cantonese: The following shows that it is bad to have a demonstrative in an existential sentence.

ii. (*jau5) ni1 bun2 syu1 nei5 jing1-goi1 tai2
have DEM CL book you ought-to read
"There's this book that you ought to read."

If definite articles occupy D, as is standardly assumed, demonstratives should not be base-generated in D. Recent analyses of the demonstrative have converged on the position that they are base-generated in the specifier position of a lower functional projection, and they may move up to D or Spec-DP (Bernstein 1997, Burgè 1996, Giusti 1997). The Cantonese demonstratives "*ni1*(this)" and "*go2*(that)", as pointed out in Cheng and Sybesma (to appear), are locatives, meaning *here* and *there* respectively. Under this analysis, they pattern not so much with demonstratives in the European languages (e.g. *this/that* in English) but more with locatives (e.g. *here/there* in English) in other languages. According to Bernstein (1997), these locatives are "reinforcers" which may well be base-generated as the head of a functional projection (She calls it *FP - the Focus phrase* - within DP[11]). Whether Cantonese demonstratives are indeed locatives is not immediately clear to me.[12] In any case, I think it is possible for Cantonese demonstratives to occupy a head position: The specifier analysis (in the above-mentioned works) assumes that demonstratives move to a higher position, but it is unclear if the Cantonese demonstratives ever move. In addition, Cantonese demonstratives cannot be used alone (as in English - see fn.12), indicating that they behave more like a functional head than a "phrasal" element in specifier position. Later, I will assume with Cheng (1998) and Li (1999a, b) that demonstratives project a functional head Dem, which is crucial in an analysis which explains how nominal expressions in Cantonese yield (in)definite interpretations (See below).

If demonstratives are not D either, the remaining possibility is that classifiers are D in Cantonese. According to Longobardi (1994, 1996), D is a universal category which determines the *referential status* of nouns (i.e. definite/indefinite/generic) by picking out individual entities from a kind-denoting noun/NP.[13] Cheng and Sybesma (to appear)

---

[11] Other linguists (Burgé 1996, Giusti 1997) just consider demonstratives to be base-generated in a lower projection without identifying it as FocusP.

[12] Actually we have to insert some "classifier-like elements" after the demonstrative in order to carry the meaning "here/there":

i. ni1 dou6
   this place
   "this place" (= here)

ii. go2 dou6
   that place
   "that place" (= there)

[13] It has been standardly assumed since Carlson (1977) that bare nouns without determiners denote kinds.

point out that the function of determiners in marking the referential status of nouns is fulfilled by *classifiers (CL)* in Cantonese. Let's look at the following pair of examples in Cantonese:

(8) keoi5 zung1 ji3 tai2 **hei3**
He/she love watch movie
'He/she loves watching movies.'
(There are no modifiers before the noun "*hei3*(movie)"; the nominal expression refers to "movies" in general, not a particular one.)

(9) keoi5 tai2 zo2 **tou3** **hei3**,
he/she watch ASP CL movie
'He/she has watched a/the movie.'
(The classifier "*tou3*" appears before the noun "*hei3* (movie)"; the nominal expression refers to a particular movie.)

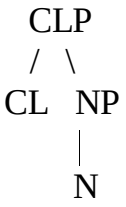
While the bare noun "*hei3* (movie)" in (8) is interpreted as generic, adding a classifier (i.e. "*tou3 hei3* (CL+ movie)") in (9) yields an existential (definite or indefinite) interpretation.

Note that the generic bare noun - "*hei3* (movie)" - is also an argument in object position. As already mentioned above in section 1, Cheng and Sybesma (to appear) consider that this bare noun has moved to an empty CL position. Sometimes, a bare noun may also give a definite reading (e.g. (10a)). Cheng and Sybesma (to appear) consider that the bare nouns in these cases also move to CL at LF (i.e. (10b)).

(10) a. keoi5 bei2 lou5 baan2 laau6.
he/she PASS boss scold
'He/she has been scolded by **the** boss.'
b. $[_{CLP}$ lou5 baan2$_i$ $[_{NP}$ t$_i$]]

Cheng and Sybesma conclude that there is at least one functional layer above Cantonese NP's:

(11)



(Cheng and Sybesma, to appear: 8, (16))

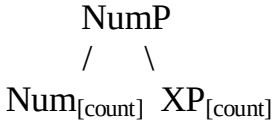
As operators binding NP's, classifiers play a function in Cantonese similar to articles in other languages. This similarity between English articles and Cantonese classifiers can be captured by saying that they both share a [referential] feature in the sense of Longobardi (1994):

(12) a. Articles (D-element in English): [referential]
b. Classifiers (D-element in Cantonese): [referential]

4 Classifiers encode a [count] feature

In constrast with English, number is not grammaticalized in nouns. Cheng and Sybesma argue that classifiers are involved in grammatical number. Adopting Doetjes' ideas (1996, as quoted in Cheng and Sybesma, to appear), Cheng and Sybesma (to appear) suggest that numerals are licensed if a nominal expression has some "syntactically visible" device which reflects the "semantic partitioning" of the noun. In English, such a device may well be number morphology. In Chinese, such a device may well be classifiers. This suggestion is in line with the long-standing intuition that classifiers somehow enable nouns to be counted, an idea which is reflected in the Chinese name for classifiers - "liang ci(measure words)".

Cheng and Sybesma (to appear) do not spell out exactly how classifiers express grammatical number. However, building on their insights (and also Doetjes' (1996, as quoted in Cheng and Sybesma, to appear), let's assume that a numeral has to combine with a nominal expression which has the "syntactically visible" feature of [count]. Let's say that this [count] feature agrees with a [count] feature of the numeral:

(13)


(XP=NP or extended projection of NP)

In English, number morphology performs the function of making the "semantic partitioning of nouns" visible in syntax. Let's say that in English count nouns encode a [count] feature but mass nouns do not. Consequently, a count noun licenses a numeral (e.g. (14a)), but a mass noun does not (e.g. (14b)).

(14) a. one/two/three/(etc.) apple(s)
[count] [count]

b. *one/two/three/(etc.) water[14]
[count]

In Chinese, however, a numeral cannot combine with an NP alone, even if the head noun refers to a countable object.[15] A classifier has to precede a noun in order to yield a grammatical sequence (e.g. (15)).[16] [17]

(15) jat1/ loeng2/ saam3/(etc.) *(go3) ping4 gwo2
one/ two/ three CL apple

To account for (15), we may say the [count] feature is not syntactically encoded in nouns (despite a conceptual substance-object distinction) but is encoded in classifiers. This explains why classifiers can turn "mass nouns" into countable expressions (e.g. (16)):

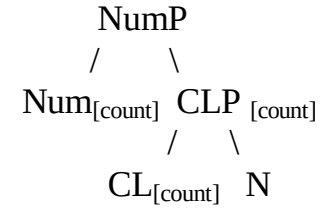
[14] Where a typically mass noun is occasionally used as a count noun, as in "two waters" or "three coffees", I assume that it bears a [count] feature.

[15] Notice that Cantonese infants have developed the ontological distinction of *substance* vs *object* despite the lack of morpho-syntactic device to express it in Cantonese (Teng 1997).

[16] There is only one special situation where classifiers are absent, that is, in *proverbs*, for instance,
i. jat1 jat6 jat1 ping4 gwo2, ji1 sang1 jyun5 lei4 ngo5
one day one apple doctors far apartme
(Literally) "An apple a day keeps the doctor away."
Nevertheless, such proverbs can be treated as "listed", being inserted as a "whole" in phrase structure (Jackendoff 1997).

[17] Contrary to all other classifiers, "di1", a classifier which refers to an indefinite amount of something, cannot be preceded by a numeral *more than one*:
i. jat1/ *loeng5/ *saam1 di1 syu1
one/ two/ three CL book
The idiosyncratic behaviour of "*di1*" - namely, that it cannot take a numeral larger than "*jat1*(one)" - may well be due to its semantics. "*Di1*" means "a small quantity of something"; in other words, "*di1*" refers to an *indefinite amount*, which resists numeral quantification. In English, we also have expressions like "*a bit of*" or "*a great deal of*", which denotes indefinite amounts. We cannot say "**two bits of*" or "**three great deals of*".

(16)



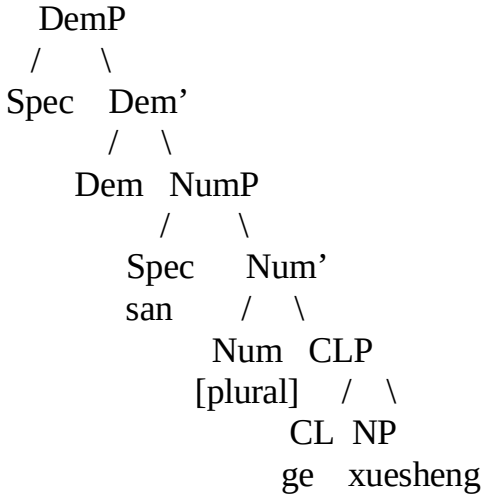
e.g. saam1 bui1 seoi2
three CL water
"three glasses of water"

On the other hand, we can also explain why classifiers are needed for count nouns after a numeral - nouns in Cantonese do not encode a syntactic feature of [count] and a numeral combines with a nominal projection which has a "syntactically visible" [count] feature. Again, this essentially follows the insights of Doetjes, Cheng and Sybesma (op. cit.).

We may now assume that classifiers have a feature [count] whereas English articles do not.

(17) Classifiers: [referential, count]

Li (1999a, b) has a different conception of how number is grammatically represented in Chinese. She suggests that the Num head carries a [singular] or [plural] feature, whereas the numeral is base-generated in Spec-NumP. For instance, a nominal expression "san ge xuesheng(three students)" has the following structure.

(18)


san ge xuesheng
three CL student
"three students"

The chief motivation behind the analysis in (18) is to capture the distribution of the plural marker "men", which Li (1999a, b) regards as realizing the [plural] feature. By positing it to be in Num, the analysis in (18) captures the fact that this marker can move up and adjoin to pronouns or proper names, which are supposed to be base-generated in D(em) (e.g. (19)). On the other hand, some "bare" common nouns may be adjoin to "men" on its way to D(em), provided that there are no intervening heads blocking the movement (i.e. CL) (e.g. (20)).

(19) a. ta-men (pronoun + men)
he/she-PL
'they/them'
b. Xiao-qiang-men (proper name + men)
[name]-PL
'Xiao-qiang and his associates'

(20) a. tungzhi-men (noun + men)
comrade-PL
'comrades'
b. *san ge tungzhi-men (Num + CL + noun + men)
three CL comrade-PL
'three comrades'

As for Cantonese, the plural marker which roughly corresponds to "men" is "dei6". Contrary to "men", "dei6" can occur with pronouns only, indicating that the latter is not a "full-fledged" plural marker. It is thus doubtful if the analysis in (18) can apply to Cantonese as well (e.g. (21))

(21) a. keoi5 dei6 (pronoun + dei6)
he/she PL
'they/them'

b. (?)Siu2-ming4 dei6 (proper name + dei6) [18]
[name] PL
'Siu-ming and his associates'

c. *lou5 baan2 dei6 (common name + dei6)
boss PL
'the boss and his/her associates'

As for Mandarin Chinese, it is unclear whether grammatical number is indeed expressed by the [singular/plural] feature as in Li's analysis (1999a, b): The problem is that a "plural" interpretation can be yielded without the marker "men" (e.g. "san ge xuesheng(three students)" in (18)).

5 Demonstratives encode a [definite] feature

The picture of how (in)definiteness is interpreted in Cantonese appears complicated. I argue against Cheng (1998) and Cheng and Sybesma (to appear) who treat a bare CLP (i.e. without an overt numeral/demonstrative) as inherently definite and a bare NumP (i.e. without an overt demonstrative) as inherently indefinite. Instead, I opt for an analysis in which definiteness is encoded in demonstratives, as suggested by Li (1998, 1999a, b).

First of all, note again that a CLP can be interpreted as definite or indefinite depending on various syntactic or discourse factors;[19] for instance, example (9), which is repeated below:

---

[18] I have observed that several speakers use this pattern where a proper name precedes "*dei6*". However, most speakers intuitively find this pattern unnatural and perfer the following one:

i. Siu2-ming4 keoi5 dei6
[name] he/she PL
"Siu-ming and his associates."

(22) keoi5 tai2 zo2 **tou3 hei3**,
he/she watch ASP CL movie
'He/she has watched a/the movie.'

In this example, "*tou3 hei3*(movie)" is interpreted as definite if the movie has been identified in the discourse. On other occasions, it may be interpreted as indefinite, if the hearer does not know in advance which movie the speaker is referring to. Extending Longobardi's (1994) idea that an indefinite interpretation is yielded when there is an empty operator position before N, Cheng and Sybesma (to appear) propose that the indefinite reading of "*tou3 hei3*(CL + movie)" is yielded with an empty head Num(=number) dominating the CLP. The definite reading is yielded when Num is absent.

That the CLP ("*tou3 hei3*(CL + movie)"), when interpreted to be indefinite, is bound by an empty Num position is further supported by the following example, where the insertion of an overt numeral forces an indefinite reading:

(23) keoi5 tai2 zo2 **jat1 tou3 hei3**,
he/she watch ASP one CL movie
'He/she has watched a movie.'

Second, notice that a classifier phrase "*bun2 syu1*(the book)" in subject position must be interpreted as definite:

---

[19] Factors affecting the interpretation of (in)definiteness include *word order* (see text below), *aspect* (e.g. (i)), and, as pointed out by Cheng (1998), *the presence of sentence-final particles* (e.g. (ii)). There may well be other factors which are still open to further research.

i. ngo4 wan2 faan1 bun2 syu1
I find ASP CL book
'I've found the book.'

ii. ngo5 wan2 dou2 bun2 syu1 laa3
I find ASP CL book C
'I've found the book.'

In (i), the aspect marker "*faan1*", literally "back", implies that the sequence "*bun2 syu1*(CL + book)" refers to "a book that I've lost", which is supposed to be known to the hearer, hence the definite reading. In (ii), the sentence-final particle "*laa1*" implies that the event is related to some prior knowledge of the listener, hence the definite reading. What distinguishes (ii) from (i) is that in (ii) the context is wider: The speaker may have lost a book but he has found it again (as in (i)), or he may have been looking for a particular book and he has found it, etc.

(24) **bun2** **syu1** m4 gin3 zo2
CL book NEG see ASP
'The book is lost.'

The absence of an indefinite reading in (24) is expected in Cheng and Sybesma's account: An empty Num head, which contributes to the indefinite reading, is not licensed in subject position because of the ECP. Taking the CLP in object position into account (i.e. (22)), it is clear that a CLP on its own (i.e. without a dominating, empty Num head) must be definite. On the other hand, a NumP (with or without overt numerals) is indefinite.[20] Since we do not expect an NP to be [±definite] (unless it is moved to D/CL), it appears that CL is [+definite] and Num is [-definite].

Nonetheless, classifiers cannot encode a fixed [+definite] value for obvious reasons: a NumP which contains a CLP is invariably indefinite.[21] Assuming that the features of functional heads are percolated to the highest node of an extended projection (Grimshaw 1991), a phrase which is indefinite but contains a [+definite] feature must crash. To illustrate, let's look at a well-formed sentence, in which the NumP is indefinite. Despite the well-formedness of (25a), we should not have a structure like (25b):

(25) a. ngo5 sau1 dou2 saam1 fung1 seon3
I receive ASP three CL letter
'I received three letters.'
(By saying "*sam1 fung1 seon3*(three letters)", I do not refer to which three letters I received, hence indefinite.)

b.
*NumP**[-definite]**
/ \
Num CLP
saam1 / \
CL**[+definite]** NP
fung1 **seon3**

[20]Cheng and Sybesma (to appear) attribute this fact to numerals being inherently *existential operators (∃)*. On the other hand, classifiers already imply an existential interpretation. Hence, it is unclear whether numerals really act as existential operators.

[21] Except the "quantity denoting expressions" which are not existential as expounded in Li (1998). See text below.

Similarly, although the NumP in (25a) above must not be definite, it is equally improbable that the numeral has a fixed [-definite] value. This is because we may have a DemP containing a NumP which is invariably definite. For instance,

(26) a. nei5tai2 zo2 **go2** gei2 fung1 seon3 mei6
you read ASP DEM several CL letter Q
'Have you read those several letters?'

b. *DemP$_{\textbf{[+definite]}}$
/ \
Dem NumP
go3 / \
Num$_{\textbf{[-definite]}}$ CLP
gei2 / \
CL N
fung1 **seon3**

In the light of this, we have to conclude that Num and CL do not encode a fixed [definiteness] value, although it appears that an overt Dem is invariably [+definite] (i.e. Nominal projections with a demonstrative are invariably definite). On the other hand, it is obvious that NumP's and CLP's in Cantonese are not random in definiteness; speakers are perfectly able to interpret the definiteness of NumP's and CLP's in different contexts (e.g. in terms of word order, predicates, sentence final particles, etc. - see above and fn.17). The question remains how CLP's and NumP's in Cantonese get the relevant (in)definite interpretations.

The answer to this question, I think, can be found by adopting A. Li's (1998) proposal on the difference between *individual denoting expressions* and *quantity denoting expressions*. Recall that NumP's generally cannot appear in subject/topic positions in Mandarin Chinese:

(27) *san ge xuesheng zai xuexiao shoushang le
three CL student P school hurt PRT
'Three students were hurt at school.'

(Li 1998: 694, (1))

However, it has also been noticed that in certain restricted contexts NumP may appear in subject/topic positions:

(28) wu ge xiaohai chi bu wan shi wan fan
five CL child eat NEG finish t en bowl rice
'Five children cannot finish ten bowls of rice.'

(Li 1998: 695, (9))

The contrast between (27) and (28) carries over to Cantonese:

(29) *saam1 go3 hok6 sang1 hai2 hok6 haau6 sau4 soeng1 laa3
three CL student P school hurt PRT
'Three students were hurt at school.'

(from Li 1998: 694, (1))

(30) m5 go3 sai3 lou6 sik6 m4 jyun4 sap6 wun2 faan6
five CL child eat NEG finish ten bowl rice
'Five children cannot finish ten bowls of rice.'

(from Li 1998: 695, (9))

Li (1998) notes that the NumP in (27) (also (29)) is interpreted as a referential argument[22] whereas the NumP in (28) (also (30)) refers to the quantity without being referential. She posits that a NumP with a referential interpretation is bound by an empty D, but "quantity denoting expressions" like the NumP in (28) and (30) are not. Accordingly, a NumP with a referential interpretation (e.g. (27), (29)) is ruled out in subject/topic positions by the *Empty Category Principle*. On the other hand, a "quantity-denoting" NumP can appear in subject/topic position without violating the ECP because there is no unbound empty operator (namely, D).

(31) a. [$_{DP}$ D[$_{NumP}$ san ge xuesheng]]
three CL students
'three students'
(individual denoting expression)

[22] Li (1998) does not explain why an existential/referential interpretion is forced here. I assume that this may have to do with the sentence-final particle (i.e. "*le*")which binds the event and gives it a "perfective" interpretation (see fn.19 above).

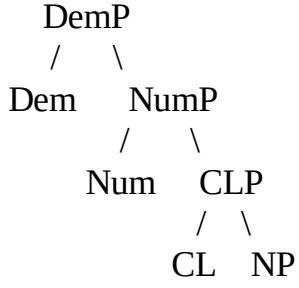
b. [$_{NumP}$ wu ge xiaohai]]
five CL child
'five children'
(quantity denoting expression)

(Refer to (27) and (28), from Li 1998: 696, (13))

Li (1998) has further illustrated the claim that "individual denoting expressions" and "quantity denoting expressions" do have different properties with regard to scope and binding. If we adopt Li's proposal, *the locus of definiteness lies in D, which is actually the position of demonstratives in Chinese*. On the one hand, this proposal captures the obvious fact that overt demonstratives precede NumP's (e.g. (26a)) and yield a [+definite] interpretation. On the other hand, as we have seen above, it is problematic to consider demonstratives to be D's - According to current theory (Longobardi 1994, etc.), D turns an NP into a referential argument, but demonstratives themselves cannot combine with nouns alone without classifiers (cf. (6)).[23] Let's say that classifiers are indeed D's as conceived in current theory, and retain the category label Dem for demonstratives.

Recasting Li's proposal again, we reach the following generalizations for Chinese:

(32) *Structure of Chinese nominal expressions*
Nominal expressions which are [referential] obligatorily project a DemP:

.


(33) *Definiteness in Chinese*
a. A definite nominal expression is yielded if Dem is overtly filled (e.g. (26a)).
b. An indefinite nominal expression is yielded if Dem is empty (e.g. (25a)).

[23] It is clear to me that classifiers are obligatory with a demonstrative in Cantonese. It is unclear whether Mandarin is the same. Tang (1990) assumes so but A. Li (1999b, fn.25) does not.

6 A microparameter: CL-to-Dem movement in Cantonese and N-to-Dem movement in Mandarin Chinese

Assuming (33), let's go back to the question of how nominal expressions yield (in)definite interpretations in Cantonese. Nominal expressions with overt demonstratives are invariably definite according to (33a) (e.g. (26a)). Nominal expressions with overt numerals (but without demonstratives) are invariably indefinite, which, according to (33a), is due to an empty Dem dominating the NumP (refer to the structure in (32)). There is actually a syntactic option by which a nominal expression with a numeral gets a definite reading, namely, by moving the numeral to Dem, but this possibility is not supported by the facts of Chinese (i.e. Nominal expression with numerals only are indefinite). I remain open as to the nature of the constraint (which may well be semantic) which bars such movement.[24] As for nominal expressions with overt classifiers only, I suggest that they yield a definite reading by *moving the classifier to Dem*, as in examples (22) and (24).[25]

The existence of such a CL-to-Dem movement is further supported by the fact that an intervening Num blocks the movement - As a result, nominal expressions with overt numerals (but without overt demonstratives) are indefinite. (Note that numerals do not move to Dem for independent reasons.) This falls out from the Head Movement Constraint. The function of numerals as blockers of movement explain why the numeral "*jat1*(one)" is inserted, even though it does not need to be overt for a CLP to yield the interpretation of "one" (e.g. (22) and (23)).

On the other hand, Mandarin Chinese does not seem to allow such movement. This is illustrated by the impossibility of nominal expressions with only classifiers in subject/topic positions:

[24] It seems that a numeral cannot move to D in other languages either (e.g. English).

[25] The overt movement of CL to Dem does *not* result in an order in which the moved classifier precedes the adjuncts of CLP. For one thing, adjuncts of CLP's are extremely restricted, for instance:

i.	keoi5	song3	zo2	*(jat1)	**daai6**	soeng1	syu1	bei2	ngo5
	he/she	give	ASP	one	big	CL	book	to	me

"He gave me a big box of books.'

Notice that the adjunct has to appear with a numeral, thus the classifier cannot move anyway due to the Head Movement Constraint.

(34) ***ben** **shu** bu hao
CL book NEG good
'The/A book is not good.'

(Cheng and Sybesma, to appear: 15, (28))

As the Mandarin classifier "*ben*" does not move up to Dem, an empty Dem category is not lexically governed, hence violating the ECP. In object positions, we would expect "*ben shu*" to appear, as Dem is governed. However, contrary to corresponding expressions in Cantonese, it cannot yield a definite reading since the classifier has to remain in CL. This prediction is again borne out:

(35) wo xiang kan **ben** **shu**.
I would like read CL book
I would like to read a book/*the book.'

(Cheng and Sybesma, to appear: 15, (27a))

Whereas Mandarin Chinese does not allow CL-to-Dem movement, it appears to be parametrized for N-to-Dem movement[26], an option not available to Cantonese in principle.[27] This claim predicts that Mandarin bare nouns can appear in subject position (via N-to-Dem movement) with a definite reading only (Chao 1968: 308, Li and Thompson 1981: 86).[28] This is borne out by the following example:

(36) **miaoer** zai shuijiao
cat PROG sleep
'The cat is sleeping.'

Cantonese bare nouns cannot appear in subject positions - Given that N cannot move up to Dem, bare nouns would violate the ECP in subject position:

---

[26] The possibility of N-to-D(em) movement for Mandarin Chinese has already been suggested by Li (1999a, b).

[27] Recall that a bare noun in Cantonese may have a definite reading (see example (10)), but, as noted by Cheng and Sybesma (to appear), these cases are restricted to some nouns referring to human beings only. Under the present analysis, these Cantonese nouns move to Dem.

[28] A Mandarin bare noun can also be generic (Li and Thompson 1981: 86). I assume that the generic bare nouns also move to Dem from N. In any case, the indefinite reading is out for a Mandarin bare noun in subject position.

(37) ***maau1** fan3 gan2 gaau3
cat sleep ASP
'The cat is sleeping.'

To turn (37) into a grammatical sequence, one has to insert a classifier, which, under the present theory, moves to Dem:

(38) **zek3 maau1** fan3 gan2 gaau3
CL cat sleep ASP
'The cat is sleeping.'

On the other hand, if N can move to Dem in Mandarin Chinese, we anticipate that a bare noun in object position can have a definite or an indefinite reading: The former is yielded by the bare noun moving to Dem; the latter is yielded by the bare noun remaining in situ. This prediction is again borne out:

(39) a. Ta zuotian mei shang **ke**
he/she yesterday not attend class
'He/she didn't attend a class/classes yesterday.'
b. Hufei he-wan-le **tang**
[name] drink-finish-ASP soup
'Hufei finished the soup.'

(Cheng and Sybesma, to appear: 1, (1b))

On the other hand, since a bare noun has to remain in situ in Cantonese, it has to be interpreted as indefinite in object position, leaving an empty Dem node. This is illustrated by the following example:

(40) ngo5 zou6 saai3 **gong1 fo3** laa3
I do ASP assignment PRT
'I've finished my assignment!'

("*gong1 fo3*(homework) does not refer to particular assignments, hence indefinite.)

The following summarizes the interpretative possibilities of bare nouns/CLP's in Cantonese and Mandarin Chinese:

(41)

	Mandarin Chinese	*Cantonese*
Bare nouns in subject position	+definite	*impossible
Bare nouns in object position	±definite	-definite
Bare CLP's in subject position	*impossible	+definite
Bare CLP's in object position	-definite	±definite

In sum, the differences of interpretative possibilities of bare nouns/CLP's in Mandarin Chinese and Cantonese fall out neatly from the postulation of CL-to-Dem movement in Cantonese and N-to-Dem movement in Mandarin Chinese. This theory resorts to syntactic mechanisms only and reduces the "Definiteness Constraint" for Chinese (Xu 1995) to the ECP. The analysis of Cheng and Sybesma (to appear) also accounts for the same range of data, but they have to invoke both syntactic and semantic mechanisms: The indefinite nominal expressions (i.e. bare nouns/CLP's) are supposed to have an empty node higher than N (i.e. CL and/or Num) whereas the definite ones are analysed as having no higher projections. A problem thus arises regarding the definite bare noun in Mandarin Chinese: Under Cheng and Sybesma's (to appear) analysis, the definite bare nouns in Mandarin only project to NP, contrary to their assumption that there is always an empty CL node binding "bare nouns" in Chinese. The solution Cheng and Sybesma suggest is that a semantic *i*-operator binds the definite bare noun in Mandarin Chinese.

7 Summary

In this paper, I have looked into the superstructure of nominal expressions in Cantonese. In particular, I have argued that a combination of the proposals by Cheng (1998), Cheng and Sybesma (to appear) and Li (1999a, b) best captures the Cantonese facts: Firstly, classifiers, rather than demonstratives, are instanstiations of D, given that only the former can bind an NP alone. Secondly, classifiers encode a [count] feature which agree with a matching [count] feature in Num. Thirdly, neither classifiers nor numerals in Cantonese carry a [definite] feature. A definite expression is yielded by lexicalizing the Dem position - In the case of Cantonese, this is accomplished by either inserting an overt demonstrative or moving a classifier to the Dem position. Looking at Cantonese in a typological context, we find that D may be instantiated not only by articles (as in English, French, etc.) but also by classifiers which encode different features (e.g.

Cantonese, and possibly other classifier languages as well). This idea is further corroborated by recent suggestions that case particles are also instantiations of D (Giusti 1995, Osawa 1997).

References

Abney, S. (1987). *The English Noun Phrase and its Sentential Aspect*. Doctoral Dissertation, MIT.

Bernstein, J.B. (1997). Demonstratives and Reinforcers in Romance and Germanic languages. *Lingua* 102, 87-113.

Burgè, L. (1996). Demonstrative movement in Spanish: A comparative approach. *University of Venice - Working Papers in Linguistics* 6(1): 1-53.

Carlson, G. (1977). *Reference to Kinds in English.* Doctoral Dissertation. University of Massachussetts, Amherst.

Chao, Y-R. (1968). *A Grammar of Spoken Chinese*. Berkeley and Los Angeles: University of California Press.

Cheng, L. L-S. (1998). *Marking Modification in Cantonese and Mandarin.* Paper presented at the Departmental Seminar. SOAS, London.

Cheng, L. L-S. and Sybesma, R. (to appear). Bare and Not-so-bare nouns and the Structure of NP. To appear in *Linguistic Inquiry*.

Chomsky, N. (1995a). *The Minimalist Program.* Cambridge, Massachusetts: MIT Press.

Chomsky, N. (1995b). Bare phrase structure. In G. Webelhuth (ed.). *The Government and Binding Theory and the Minimalist Program*. 383-439. Oxford: Blackwell.

Cinque, G. (1994). On the evidence of partial N-movement in the Romance DP. In G. Cinque, et al. (eds.). *Paths towards Universal grammar: Studies in Honor of Richard S. Kayne*. 85-110. Washington DC: Georgetown University Press.

Ernst, T. (1991). A phrase structure for tertiaries. *Syntax and Semantics 25*, 189-207.

Gao, Q. (1994). Chinese NP structure. *Linguistics* 32, 474-510.

Giusti, G. (1995). A unified structural representation of (abstract) case and article. In H. Haider et al. (eds.). *Studies in Comparative Germanic Syntax,* 77-93.

Giusti, G. (1997). The categorial status of determiners. In L. Haegeman (ed.). *The New Comparative Syntax*. 95-123. Addison Wesley Longman.

Grimshaw, J. (1991). *Extended projections.* Manuscript.

Haegeman, L. and Guéron, J. (1999). *English Grammar*. Oxford: Blackwell.

Higginbotham, J. (1985). On semantics. *Linguistic Inquiry* 16, 547-593.

Jackendoff, R. (1997). *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.

Kayne, R. (1994). *The Antisymmetry of Syntax.* Cambridge, Massachusetts: MIT Press.

Li, C. and Thompson, S. (1981). *A Functional Reference Grammar of Mandarin Chinese.* Berkeley: University of California Press.

Li, Y-H. A. (1998). Argument determiner phrases and number phrases. *Linguistic Inquiry* 29(4), 693-702.

Li, Y-H. A. (1999a). *Form-Function Correspondence: Structures and Interpretations of Nominal Expressions in Mandarin Chinese.* Paper presented at Departmental Seminar, SOAS, London.

Li, Y-H. A (1999b). Plurality in a classifier language. *Journal of East Asian Linguistics* 8, 75-99.

Lin, J-W. (1997). Noun phrase structure in Mandarin Chinese: DP or NP? *Chinese Languages and Linguistics* 3, 401-434.

Longobardi, G. (1994). Reference and proper names: A theory of N movement in syntax and logical form. *Linguistic Inquiry* 25(4), 609-665.

Longobardi, G. (1996). *The Syntax of N-raising: A Minimalist Theory*. OTS Working Papers. Research Institute for Language and Speech, Utrecht University.

Matthews, S. and Yip, V. (1994). *Cantonese: A Comprehensive Grammar*. London: Routledge.

Osawa, F. (1998). The emergence of the D system and the demise of morphological case in English. *UCL Working Papers in Linguistics* 10, 467-488.

Progovac, L. (1998). Determiner phrase in a language without determiners. *Journal of Linguistics* 34, 165-179.

Quirk, R., Greenbaum, S., Leech, G. and Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.

Ritter, E. (1995). On the syntactic category of pronouns and agreement. *Natural Language and Linguistic Theory* 13, 405-443.

Stowell, T. (1989). Subjects, Specifiers and X-bar theory. In M. R. Baltin and A. S. Kroch (eds.). *Alternative Conceptions in Phrase Structure*, 232-262.

Stowell, T. (1991). Determiners in NP and DP. In D. Bouchard and K. Leffel (eds.). *Views on Phrase Structure*, 37-56. Kluwer Academic Publishers.

Szabolcsi, A. (1994). The noun phrase. In F. Kiefer and K. Kiss (eds.). *Syntax and Semantics 27: The Structure of Hungarian*, 179-274.

Tang, J. C-C. (1990). *Chinese Phrase Structure and the Extended X' Theory*. Doctoral Dissertation. Cornell University.

Teng, Anne, Y-Y. (1997). *Count/Mass Distinction in Cantonese and its Acquisition*. Manuscript. University College London.

Tse, Y-K. (1990). *A Brief Survey of the Syntax of "Europeanized" Modern Chinese* (in Chinese). Gwong- Ming Books Co. Ltd.

Xu, L-J. (1995). Definiteness effects on Chinese word order. *Cahiers de Linguistique -Asie Orientale* 24(1), 29-48.

Zwart, J-W. (1997). The Germanic SOV languages and the Universal Base Hypothesis. In L. Haegeman (ed.). *The New Comparative Syntax*, 246-267. Addison Wesley Longman.