On Bare Classifier Phrases

Li Julie Jiang[1] and Suhua Hu[2]

[1]*Harvard University,* [2]*Minzu University of China and Harvard-Yeching Institute*

This paper investigates bare classifier phrases (ClPs), the phrase consisting of only a classifier and a noun [Cl + N], in several different languages with respect to semantic interpretations and syntactic distributions. In the literature, it has been discussed that not all classifier languages allow bare CLPs (Cheng and Sybesma 1999, 2005). In those classifier languages which allow bare CLPs, the semantic interpretations and syntactic distributions of bare CLPs are quite restricted. For example, indefinite bare CLPs can only appear in object positions but not subject positions in all languages reported. In this paper, we present new data from Yi, a Sino-Tibetan language with SOV word order, which shows that bare ClPs receive indefinite interpretation and can appear in both subject and object positions. The newly discovered data cast doubt on the previous empirical generalizations and analyses on bare classifier phrases. We propose a universal structure for bare classifier phrases by introducing an Argumental Head which is the locus of different interpretations and is subject to parameterization, which is free from the empirical problems involved in previous analyses.

1. Introduction

An important property of classifier languages, such as Mandarin, Cantonese, and Japanese, is that all nouns are like mass nouns in needing a measure word or classifier to be counted by numerals. For example, Classifier Phrases (CLPs) in Mandarin take the form of 'Num+CL+Noun', as shown below.

(1) *san*   **(ge)*   *ren*   (Mandarin)
     three   CL   person
     'three persons'

(2) *sāam*   **(zek)*   *gau*   (Cantonese)
     three   CL   dog
     'three dogs'

(3) *hong*   *san*   **(satsu)*   (Japanese)
     book   three   CL
     'three books'

However, not all classifier languages allow bare classifier phrases (bare CLPs)—the phrase consisting of only a classifier and a noun without numerals, exampled in (4) and (5) (with bare CLPs underscored).

(4)   *zek gau zungji sek juk.*   (Cantonese)
        CL dog   like   eat meat
        'The dog likes to eat meat.'
        Not: 'Dogs like to eat meat.'/ 'A dog likes to eat meat.'

(5)   **jia gau be   lim zhui.*   (Min)
          CL dog want drink water
        Intended: 'The dog wants to drink water.'

                                                  (Cheng and Sybesma 1999, 2005)

In those classifier languages which allow bare CLPs, the semantic interpretations and syntactic distributions of bare CLPs are very restricted (Cheng and Sybesma 1999, 2005). Cheng and Sybesma (1999, 2005) discuss the distribution and interpretation of bare CLPs in four Chinese dialects. They propose that the classifier can be viewed as a counterpart of the determiner in Romance and Germanic languages, based on the fact that bare CLPs phrases are definite and can freely occur in argument positions in Cantonese. According to them, the head of CLPs is the locus of definiteness, the Numeral Phrase (NumP) is the recourse for indefiniteness, the Num and CL may be left empty, and the NumP head Num can undo the definiteness introduced by the CLP. Their proposal well explains the syntactic and semantic differences of bare CLPs in four Chinese dialects (Cantonese, Mandarin, Wu, and Min).

Simpson (2005) presents a head movement analysis for bare CLPs by applying Longobardi's (1994) DP hypothesis to Cantonese. Specifically Simpson treats Cantonese definite CL-NPs as a result of CL-to-Num-to-D head movement. In this system, the head of Determiner Phrases (DPs) is the locus of definiteness and indefiniteness. Specifically, when the D head is filled, the DP receives the definite interpretation; when the D head is empty, the DP gains the indefinite interpretation, and the empty D head is subject to Empty Category Principle (ECP). Simpson's approach provides a unified analysis for nominal phrases in both classifier languages without overt determiners and languages with overt determiners (Romance and Germanic languages). It can also explain the subject-object asymmetry of bare indefinite CLPs in languages such as Cantonese and Mandarin.

In this paper, we are going to introduce newly discovered data from Yi, a Sino-Tibetan language with SOV word order. We show that bare CLPs in Yi can freely occur

in argument positions, the same as those in Cantonese. But different from Cantonese, Yi bare CLPs only receive indefinite interpretations. The new data cast doubt on previous accounts for bare CLPs shown above, namely, first the head of CLPs might not be the locus of definiteness as bare CLPs can only be indefinite, and secondly the indefinite CLPs are not subject to the ECP restriction as they can freely appear in both subject and object positions.

To account for all the empirical data that we have so far, we propose an alternative account which argues that neither the head of DP nor the head of CLP is the locus of definiteness or indefiniteness and that it is the head of Argumental Phrases (ArgP) which contributes to the different semantic interpretation of nominal phrases and is subject to parameterization. This new account is expected to be free from the empirical problems involved in the previous studies. Our account has two further consequences. One is that it can help explain why other SOV languages such as Japanese and Korean do not allow bare classifier phrases. The other is that it can shed light on the structure of nominal phrases in general.

This paper is organized in the following way: Section 2 presents previous observed data on bare CLPs from a variety of classifier languages as well as newly discovered data in Yi. Section 3 shows previous analyses on bare CLPs and their problems. In Section 4, we propose an alternative account for bare CLPs in all languages we have shown. We show that our proposal can explain not only previously established data, but also the newly discovered data and language variation. Section 6 draws a conclusion.

2. Previous established data and newly discovered Data

Cheng & Sybesma (1999, 2005) discuss the interpretation and distribution of bare CLPs in four Chinese dialects. Three of them—Cantonese, Mandarin and Min well represents the distinctive difference of bare CLPs in the classifier languages they discuss. Cantonese allows bare CLPs [CL-NP] appear in both subject and object positions. When bare CLPs appear in object positions, they can either be definite or indefinite (nonspecific), as shown in (6) and (7); when bare CLPs appear in subject positions, they only receive definite reading, exemplified in (8).

Cantonese:

(6)   *ngo zungji tong <u>zek gau</u> waan*.   (Object: definite)
          I   like   with CL dog play
          'I like to play with the dog.'
          Not: 'I like to play with a dog.' / 'I like to play with dogs.'

(7)   *ngo soeng maai <u>bun syu</u> (lei   taai)*.   (Object: indefinite)
          I   want   buy   CL book come read
          'I like to buy a book (to read).'

(8)   *zek gau   zungji sek juk.*   (Subject: definite only)
        CL dog   like   eat meat
        'The dog likes to eat meat.'
        Not: 'Dogs like to eat meat.'/ 'A dog likes to eat meat.'
                                                            (Cheng and Sybesma 1999)

Mandarin only allows bare CLPs to appear in object positions with indefinite (nonspecific) reading, illustrated in (9) and (10). Bare CLPs are disallowed to appear in subject positions, as shown in (10) (Cheng & Sybesma 1999, 2005).

Mandarin:
(9)   a. *wo xiang   mai ben shu.*   (Object: indefinite)
           I   would-like buy CL book
           'I would like to buy a book.'
        b. **wo xiang gen   zhi gou   wan*   (Object: *definite)
            I   want with CL dog play
           Intended reading: 'I want to play with the dog.'

(10)   a. **zhi gou   yao guo malu.*   (*Subject)
            CL dog   want cross road
           Intended: 'A dog wants to cross the road.'
        b. **zhi gou xihuan chi rou.*
            CL dog like   eat meat
           Intended: 'A dog likes to eat meat.'
                                                            (Cheng and Sybesma 1999)

The contrast in (9) and (10) shows that Mandarin exhibits a subject-object asymmetry with respect to the positions that bare CLPs can occur in a sentence, which is not attested in Cantonese.

        Differing from Cantonese and Mandarin, Southern Min does not allow bare CLPs at all. The classifier can never occur without being preceded by either a numeral or demonstrative, showed in (11) (Cheng & Sybesma 2005).

Southern Min:
(11)   a. **ua siuN bue bun zhu*   (*Subject)
            I   want buy CL book
           Intended: 'I would like to buy a book.'
        b. **jia gau be   lim zhui*   (*Object)
            CL dog want drink water
           Intended: 'The dog wants to drink water.'   (Cheng and Sybesma 2005)

In Wu (Fuyang dialect)[1], bare CLPs appear in either preverbal positions (subject and shifted object positions) or postverbal positions (base-generated object positions). Preverbal bare CLPs in Wu (Fuyang) have a definite interpretation, as shown in (12) and postverbal bare CLPs have an indefinite interpretation, as shown in (13) (Li 2011).

Wu (Fuyang):
(12)   a. *tsəʔ giu sɿ-ŋiɔ die.*   (Subject: definite)
            Cl   dog die   Part
            'The dog died.'

        b *ŋɤ saŋ gə yoʔ bu tsʰotsɿ ma le uœlœ die.*   (Shifted Object: definite)
            I   last Cl month Cl car   buy Perf back Part
            'I went to buy the car last month.'

(13)   *ŋɤ maʔ le bu tsʰotsɿ. n tsʰœ-tsʰa-kʰan zɿ goʐ tsoʔ ?*   (Object: indefinite)
        I   buy Perf Cl car.   you guess   be what car
        'I bought a car. Can you guess what car it is?

(Li 2011)

The languages allowing bare CLPs above are all SVO languages, and it seems that SOV languages do not allow bare CLPs at all, such as Japanese or Korean, illustrated in (12) and (13) respectively.

Japanese:
(14)   a. **kodomo ri-ga   benkyoo shite-iru*   (*Bare CLPs)
            child   Cl -Nom study   do-be
            Intended reading: 'One/The child is studying.'
        b. **John-wa hong satsu-o   katta*
            John-Top book CL -Acc bought
            Intended reading: 'John bought a book.'

Korean:
(15)   a. **soi mali-ka   swuley-lul kkul-ko iss-ta*   (*Bare CLPs)
            cow Cl -Nom cart-Acc   pull-Del Prog-Decl
            Intended reading: 'The/One cow is pulling a cart'

---

b. *na-nun   haksayng myeng-ul   po-ass-ta
   I-Top   student   Cl-Acc   see-Past-Decl
   Intended reading: 'I saw one/the student.'

In this paper, we present some newly discovered data from Yi, a Sino-Tibetan language with an obligatory classifier system. The same as Japanese and Korean, Yi is a SOV language; however Yi allows bare CLPs, which can appear in both subject and object positions, with indefinite interpretations only, demonstrated in (14) and (15)

Yi:
(16)   *ke ma   ngo xi   la*.   (Subject-indefinite)
        dog CL we bite come
        'A dog came to bite us'
        Not: 'The dog came to bite us.'

(17)   *ngo ke ma vu bbo mi*   (Object-indefinite)
        we   dog CL buy go will
        'We are going to buy a dog.'
        Not: 'We are going to buy the dog.'

To summarize the data that we have shown so far, in SOV languages, neither Japanese nor Korean allows bare CLPs, but Yi allows bare CLPs in both subject and object positions with indefinite reading only. In SVO languages, Min disallows bare CLPs, Mandarin allows indefinite bare CLPs in object positions only, Wu (Fuyang) allows definite bare CLPs in subject positions and indefinite bare CLPs in object positions, and Cantonese allows definite bare CLPs in both subject and object positions and indefinite bare CLPs in object positions. The distribution and the interpretations of the bare CLPs in these seven languages are summarized in (16).
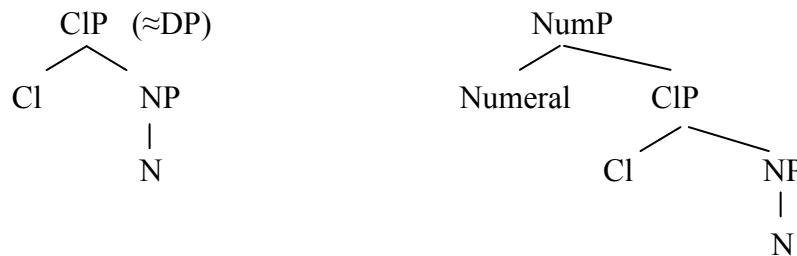
(18)

	Bare Classifier Phrases						
	Verb Medial (SVO) languages				Verb Final (SOV) languages		
	Mandarin	Min	Cantonese	Wu	Yi	Japanese	Korean
Def	*	*	Subject/Object	Subject	*	*	*
Indef	Object	*	Object	Object	Subject/Object	*	*

3. Previous Analyses

on the fact that in Cantonese bare CLPs are definite and can freely occur in argument positions, Cheng & Sybesma (1999) propose that classifiers are like determiner in Romance and Germanic languages which turns predicates to arguments and yields the definite interpretation (comparable to an iota operator 'ι'). They propose that all definites ([CL-NP]s and bare nouns) and generics have the structure in which a CL is filled either by the ι operator (realized as an overt classifier) or a moved N, as in illustrated in (19a). For the indefinites (Num-C-NPs, Cl-NPs, and bare nouns), Cheng and Sybesma propose that they all have the structure in (19b) (Num and CL may be left empty), in which the head of NumP can undo the definiteness introduced by the head of ClP.
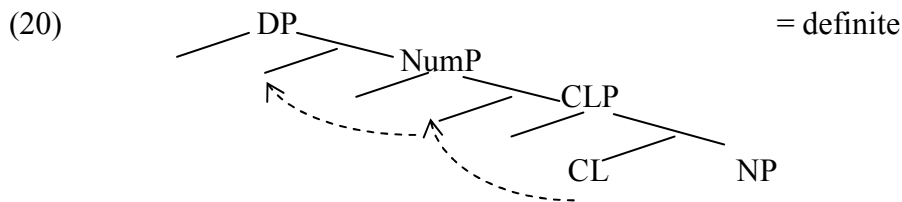
(19)   a. Definite NPs ([Cl-NP], N)   b. Indefinite NPs ([Num-Cl-NP], [Cl-NP], N)
            Generic NPs (N)

```
        ClP  (≈DP)                              NumP
       /    \                                  /     \
     Cl      NP                          Numeral      ClP
             |                                        /   \
             N                                      Cl     NP
                                                           |
                                                           N
```
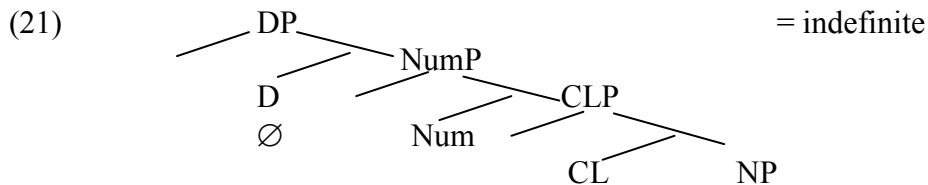
Cheng and Sybesma explain the interpretational and distributional differences of bare CLPs phrases by following some of the ideas developed by Longobardi (1994). When the CL position is filled (by a classifier or a moved N), the CLP receives either a definite or generic interpretation and is not limited to occur in lexically governed positions. When the projection above NP involves an empty head, it must be lexically governed, and this explains why the indefinites bare CLPs are distributionally restricted to lexically governed positions. Specifically, they propose that Cantonese differs from Mandarin and Min in whether or not it is possible to have an overt classifier without a numeral and whether definiteness is expressed by a segmental operator ι in Cl in the form of a full-fledged classifier. Min differs from Cantonese and Mandarin in that Min cannot have empty numerals while other classifier languages can Cheng and Sybesma.

However, their analysis has the following problems. First, the head of CLP might not be the locus of definiteness as they propose. Their argument for this proposal is based on the fact that bare CLPs in Cantonese are definite and can freely occur in argument positions. However, bare CLPs in Yi can also free appear in argument positions, but they only receive indefinite readings. Secondly, the government-based account cannot explain why indefinite bare CLPs in Yi can freely appear in both subject and object positions, we will elaborate this point after we introduce Simpson's analysis as the same problem also arise for him.

Simpson (2005) argues against the view that CLs correspond to the definite article since definite articles are always higher than Numeral Phrases in languages with overt determiners, but CLs occur lower than numerals. Simpson proposes a head movement analysis for bare CLPs by applying Longobardi's (1994) DP hypothesis to Cantonese. Longobardi's hypothesis has several important components. The first one is that an empty D head leads to a default existential interpretation for the DP; secondly, an empty head must be lexically governed (as a result of ECP constrain); thirdly, if the D head is filled, the DP receives a definite interpretation. Simpson analyzes definite CL-NPs in Cantonese as CL-to-Num-to D head movement, as shown in (20).

(20) DP   = definite

NumP

CLP

CL   NP

When the head of CLP undergoes CL-to-Num-to-D movement, the D head position is filled, and the whole bare CLP receives a definite interpretation—this is the case for the definite bare CLPs in Cantonese. When no movement occurs from CL to D position, the D head position remains unfilled, and the DP receives a default existential reading, illustrated in (21).

(21) DP   = indefinite

NumP

D

CLP

∅

Num

CL   NP

Syntactically, when the D head is not filled, the distribution of the DP is restricted to lexically governed positions because of the ECP constraint. This explains why indefinite bare CLPs can only occur in object positions but not subject positions in the three languages: Mandarin, Cantonese, and Wu (Fuyang).

Simpson's analysis well explains the language variation with respect to different semantic interpretation of bare CLPs in these three languages and keeps a universal structure—DPs for all nominal phrases. However, there are several empirical problems for this DP Hypothesis-based analysis. The first one is that the government-based account cannot explain why the indefinite bare CLP can freely appear in both subject and object position in Yi. This is the same problem that Cheng and Sybesma's analysis has, as mentioned above. Specifically, if assuming Longobardi's DP Hypothesis as Simpson does, Yi should on the one hand have an unfilled empty D head in order to derive the indefinite reading and on the other hand have a filled D head in order to avoid the ECP
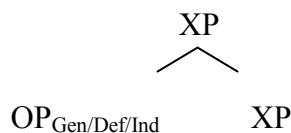
violation as bare CLPs can freely appear in argument positions. However, this is a contradiction. Secondly, if assuming head movement from CL to D position for definite readings in Cantonese, Simpson's analysis cannot explain why N cannot undergo the same type of movement to D position to get definite interpretation for bare NPs in Cantonese. Although bare CLPs can have definite readings, bare NPs in Cantonese cannot be interpreted as definite but only generic and indefinite (Cheng and Sybesma 2005). If the head of CLP undergoes head movement to D position to get a definite interpretation, one should expect that the head of NP—N should be able to move to D position to get the definite reading for bare NPs as well. However, it is not the case.

Before moving on to an alternative account to be proposed in this paper, we'd like to review the puzzles that we need to solve. The first puzzle is what the source of definite and indefinite is—is it the D head, or the CL head, or something else? The second puzzle is the language variation with respect to the different syntactic distribution and semantics interpretations as shown in section 2. In the next section, we are going to present an alternative account for these puzzles.

4. An alternative account

Based on the fact that as long as an expression could denote definite, indefinite, or generic, it can serve as an argument (no matter what the 'label' of that expression has), we propose an Argumental Operator Hypothesis, which says that as long as an argumental operator merges with a phrase, it will make that phrase argumental. Semantically, the Argumental Operator takes a type <e, t> denoting property and returns a type <e> entity. Specifically, the argumental operator can apply at any level—bare NP level, CLP level, NumP level. And there are three types of argumental operators—genetic, definite, and indefinite, as shown below.

(22)

XP

OP$_{Gen/Def/Ind}$ XP

(XP could be NP, ClP, NumP)

In (22), the XP could be bare noun phrases, bare classifier phrases, or numeral phrases. We further assume that languages differ in choosing different types[2] of operators to apply at different levels. Now, let us show how to explain the two puzzles—the locus of different interpretations and language variation in these classifier languages.

---

[2] There could be a hierarchical semantic ranking of these three types of operators, as Chierchia (1998) and Dayal (2004) suggest.

By assuming the Argumental Operator Hypothesis, the head of classifier phrases will not be the locus of definiteness, whereas the OP is the locus of the definiteness/indefinites/genericity. A question may arise here, namely, why there isn't a language with generic bare CLPs? If we assume that all three types of argumental operators (definite, indefinite, generic) can apply at different levels, we should expect generic bare CLPs as the OP_{Gen} should be able to apply at CLP level. Indeed, we do find a language—Zhuang, a Sino-Tibetan classifier language, which allows bare CLPs in both subject and object positions with generic interpretations besides definite interpretations, exemplified in (23) and (24) respectively .

Zhuang:

(23) *[tu⁰ be⁴] saµ¹ iu²   [tu⁰ mou¹]*   (Generic)
       CL sheep clean more than CL pig
       Sheep are cleaner than pigs.

(24)   a. *[ko:ŋ¹ ha:k⁸] ʔeu¹ te¹ pai¹ ham⁸ nai⁴.*   (Definite)
          Cl   officer ask him go   night this
          'The officer asked him to go there tonight.'
       b. *pai² nai⁴ [ʔan¹ ka:ŋ¹] hi⁴ wa:i⁶.*
          then   Cl jar   also broken
          'Then, the jar is also broken.'

(Example from Liu 2010)

Next, we are going to show how our account explains the language variation with respect to bare CLPs that we have presented so far. As the three argumental operators (genetic, definite, and indefinite) can apply at any level—bare NP level, CLP level, NumP level, languages differ in choosing different types of operators to apply at different levels. For languages that do not allow bare CLPs, such as Japanese, Korean, and Min, the argumental operator simply cannot apply at bare ClP level. For languages that allow bare CLPs, they differ in choosing different types of argumental operators to apply. In both Cantonese and Wu (Fuyang), only OP_{Def} applies at the bare CLP level. In Yi both OP_{Def} and OP_{Indef} can apply: when the OP_{Indef} applies, we get the indefinite bare CLPs, and when OP_{Def} applies, we will have a definite [NP-Cl-Su] phrase in Yi (see Jiang and Hu to appear for the discussion on Su in Yi). As for the indefinite interpretation of bare CLPs in Cantonese, Wu (Fuyang), and Mandarin, we assume that an empty numeral 'one' is present in the structure, in the same line as Cheng and Sybesma (1999) and Yang (2001) propose. Consequently, the indefinite bare CLP in these three languages is not a true bare classifier phrase; instead, it is a numeral classifier phrase with an empty one [e_{one}-CL-NP].

With regard to the subject-object asymmetry of the [e_{one}-CL-NP] phrase in these three languages, we think that it is an independent issue for all indefinite nominal phrases

in Chinese types of languages as other classifier languages, such as Japanese and Korean, do not have such a semantic constraint for subjects. As Chao (1968) and Li and Thompson (1981) observe, indefinite subject are not well-suited for the subject position in Chinese, and many different explanations have been proposed for this phenomenon, such as the extended Mapping Hypothesis by Tsai (1999, 2001, 2008), the Hypothesis on Constraining the Eventuality Argument by Huang (1996: 13), and the clitic hypothesis for bare classifiers by Yang (2001:72). Here, we are not going to propose any new explanation for this indefiniteness-related subject-object asymmetry, we will treat it as non-ECP related issue but an independent issue for Chinese type of languages as other linguists assume.

5. Conclusion

In this paper, we examine bare classifier phrases (ClPs) in several different languages—Cantonese, Min, Mandarin, Wu (Fuyang), Japanese, Korean, with respect to different semantic interpretations and syntactic distributions. We introduce new data from Yi, a Sino-Tibetan language with SOV word order, in which bare ClPs only receive an indefinite interpretation and can freely appear in both subject and object positions. The newly discovered data cast doubt on the previous empirical generalizations and analyses on bare classifier phrases. We present an alternative account which is free from the empirical problems for the previous analyses as pointed in section. Specifically we propose an Argumental Operator Hypothesis, which says that as long as an argumental operator merges with a phrase, it will make that phrase argumental. Semantically, the Argumental Operator takes a type <e, t> denoting property and returns a type <e> entity. Furthermore, the argumental operator can apply at any level—bare NP level, CLP level, NumP level. And there are three types of argumental operators—genetic, definite, and indefinite. With this Argumental Operator Hypothesis, we not only have an universal structure for all classifier languages, we can also explain language variation—languages differ in choosing different types of operators to apply at different levels.

REFERENCE

Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.

Cheng, Lisa Lai-Shen, and Rint Sybesma. 1999. Bare and Not-So-Bare Nouns and the Structure of NP. *Linguistic Inquiry*, Volume 30, Number 4, P:509-542.

Cheng, Lisa Lai-Shen, and Rint Sybesma. 2005. Classifiers in four varieties of Chinese. In *The Oxford handbook of comparative syntax*, ed. by Guglielmo Cinque and Richard Kayne, 259–292. Oxford: Oxford University Press.

Chierchia, Gennaro. 1998. Reference to kinds across languages. Natural Language Semantics 6:339–405.

Dayal, VENEETA. 2004. Number Marking and (In)definiteness in Kind Terms. *Linguistics and Philosophy* 27.4, pp. 393-450.

Liu, Danqing. 2010. Yuyan Kucun Leixing Xue [linguistic inventory typology]. presented at International Symposium for Comparative and Typological Research on Languages of China, Hong Kong.

Huang, Shizhe. 1996. Quantification and Predication in Mandarin Chinese: A Case Study of DOU. Doctoral Dissertation: University of Pennsylvania.

Jiang, L. Julie and Suhua Hu. To appear. An overt determiner in a classifier language. In *Proceedings of GLOW-in-Asia VIII*.

Li, Xu ping. 2011. Doctoral dissertation, Bar-Ilan University.

Longobardi, Giuseppe. 1994. Reference and proper names: a theory of N-movement in syntax and logical form. *Linguistic Inquiry* 25:pp.609-665.

Simpson, Andrew. 2005. Classifiers and DP structure in Southeast Asia. In *The Oxford handbook of comparative syntax*, ed. by Guglielmo Cinque and Richard Kayne, 806–838. Oxford: Oxford University Press.

Tsai, W.-T. Dylan. 1999. *On Economizing the Theory of A-bar Dependencies*. New York: Garland [Originally: MIT Ph.D. Dissertation, 1994]

Tsai, W.-T. Dylan. 2001. On subject specificity and theory of syntax-semantics interface. *Journal of East Asian Linguistics* 10: 129–168.

Tsai, W.-T. Dylan. 2008. Object specificity in Chinese: A view from the *v*P periphery. *Linguistic Review* 25, 479-502.

Yang, Rong. 2001. Common nouns, classifiers, and quantification in Chinese. Doctoral dissertation, Rutgers University.