

The theory of argument formation—A view from Mandarin and Yi*

L. Julie Jiang University of Hawaii at Manoa

1 Introduction

This paper is about nominal argument formation in classifier languages (CILs) and its consequences for nominal argument formation in general. A universal property of nouns across languages is their ability to serve as the arguments of predicates. For example, in the sentence *boys are singing outside*, the noun *boys* which appears in the subject position serves as the argument of the verb *ate*. How nouns become arguments, however, has been subject to much debate. Some authors claim that nouns have fixed denotations across languages (e.g. properties) and must co-occur with an article-like determiner (D), distinct from other members of the determiner family like *that* or *this*, in order to serve as an argument (e.g. Longobardi 1994 et seq; Borer 2005). Other authors claim that languages vary in how they categorize their nouns and that whether or not nouns require Ds when serving as arguments is a language-specific property, such that some languages require Ds for all of their argumental nouns, and others do not (e.g. Chierchia 1998; Dayal 2004; Bošković 2005 et seq). Another way to conceive of this debate is as a difference in whether Determiner Phrases (DPs) are universally projected across languages. Only the latter group of authors allows for variation and parameterization in the obligatory status of D.

Until now, the debate surrounding the universal status of Ds as ‘argumentizers’ for nouns has mainly focused on data from number marking languages (NMLs), that is, languages with morphological exponents of grammatical number. For example, in English, whether or not we are talking about *one book* or *many* will determine the form of the noun we use: *book* vs. *books*. Additionally, numerals in number marking languages combine directly with nouns: *one book* vs. *two books*. Number marking languages split on whether they possess overt determiners. While languages like English and French do have overt determiners, languages like Russian and Hindi do not. These two types of languages further differ in whether they freely allow bare nouns to serve as arguments: English and French impose restrictions on the possibility of bare arguments (1), whereas in Russian and Hindi the distribution of bare nouns is unconstrained (2).

- (1) a. *I bought *(a) dog.* (English)
 b. *Dog*(s) like meat.*
 c. **(La) baleine est réputée être le plus grand mammifère* (French)
 the whale is deemed be the more grand mammal
 ‘The whale is regarded as the biggest mammal.’ (Chierchia 1998: 341, ex.3b)

* Acknowledgements to be added

Example (1) (Continued)

- d. **(Les) baleines sont en train de disparaître.*
the whales are in train of disappear
i. 'The set of sub-species of whales is becoming extinct.'
ii. 'Whales are becoming extinct.' (Vergnaud and Zubizarreta 1992, ex. 84a)
- (2) a. *Ljudi proizoshli ot obez'jan* (Russian)
Men evolved from apes.
'Men have evolved from apes.'
- b. *V komnate byli mal'chik i devocka*
in room were boy and girl
'A boy and a girl were in the room.'
- c. *kutte bahut bhau Nkte haiN* (Hindi)
dogs lot bark
'The dogs/Dogs bark a lot.'
- d. *kamre meN cuuhaa hai*
room in mouse is
'There's a mouse in the room'. (Dayal 2004, ex. 15b, 17a, 19a, b)

What both sides of the debate surrounding the universal status of Ds require is an account of these latter class of languages where we find bare nouns as arguments (e.g. (2)). Proponents of the universal DP hypothesis argue by analogy with languages like English and French, claiming that even when we find bare nouns as arguments, a covert D is present in the syntax to argumentize these nouns (e.g. Li 1998; Progovac 1998). In contrast, objectors to the universal DP hypothesis take bare nouns at face value and claim that languages that freely employ them do so without additional structure that serves the role of argumentizing nouns (e.g. Chierchia 1998; Dayal 2004; Bošković 2005 et seq).

Unfortunately, focusing only on number marking languages leaves the landscape of nominal arguments incomplete, and worse, ignores facts relevant to the universal DP hypothesis. Not all languages are number-marking: CILs like Mandarin and most other East Asian languages do not express grammatical number, and more importantly, do not allow numerals to combine directly with nouns. For example, in (3), a classifier (Cl in short) is required to connect a noun with a numeral regardless of whether the noun is conceptually count or mass.

- | | | | |
|-----|--|---|------------|
| (3) | a. <i>*san xuesheng</i>
three student | a'. <i>san ge xuesheng</i>
three Cl student
'three students' | (Mandarin) |
| | b. <i>*yi shui</i>
one water | b'. <i>yi sheng shui</i>
one Cl _{liter} water
'one liter of water' | |

In this paper, we consider how the facts from CILs fit into and inform the debate about the DP hypothesis. More generally, we use the facts from CILs in addition to those familiar from NMLs to better understand how arguments are formed across languages.

Our discussion focuses on two types of CILs. The first type, as exemplified by Mandarin, is superficially similar to Hindi and Russian in that it freely allows bare arguments and does not have any overt evidence of the syntactic category D. This type of CILs has been widely discussed in the literature (see Tang 1990; Li 1997, 1998, 1999; Cheng and Sybesma 1999; Simpson 2005; Zhang 2013 among many others). The second type, as exemplified by Yi (a Sino-Tibetan language with SOV word order), is previously unattested and in fact predicted to not exist (e.g. Chierchia 1998): this is a CIL with an overt article-like definite determiner, distinct from other members of the determiner family like *that* or *this* (as we will see in Section 2). Such a CIL also freely allows bare nouns to occur in argument position, in the same way as Mandarin, Russian, and Hindi do. The facts from Yi confirm one of Chierchia’s 1998 generalizations that *all classifier languages allow bare arguments*. Nevertheless, it does raise two main puzzles: (i) how and why does D develop in languages where it is not required? (i.e. does D work in the same way in CILs and NMLs?), and (ii) why is it the case that *no* known CIL requires overt Ds for nouns to occur in argument position, including those like Yi that have a ‘real’ article (i.e. why is there no analogue of French among the CILs?).

These two puzzles lead to a more general theoretical issue: whether or not argumentizing nouns always requires the presence of a determiner in the syntax. The discovery of a classifier language with an overt article-like determiner may at first seem to tilt the balance in favor of the universal DP hypothesis: having found overt D in one classifier language, one could argue by analogy in favor of (null) determiners everywhere. However, we use the facts from Yi to argue that the opposite is the case: the limited and predictable distribution of determiners in classifier languages in fact argues for the view that nouns have different denotations in NMLs and CILs (i.e. properties or entities) and that CILs without overt Ds like Mandarin are *D-less* in the nominal domain. Hence, the comparative analysis of the two types of classifier languages supports the view that determiners are not obligatory in argumentizing nouns.

This paper is organized as follows. Section 2 compares Mandarin with Yi.¹ The comparison between these two CILs raises a series of general puzzles concerning ways of marking definiteness, behaviors of demonstratives, and the status of D. For instance, demonstratives and the overt article-like determiner *su* in Yi appear to combine only with higher nominal projections (i.e. projections above bare nouns) and can never combine with bare nouns (4). In contrast with this, demonstratives in Mandarin can freely combine with higher nominal projections as well as bare nouns (5).

¹ Most data from Yi in this paper are based on our consultant work with a native Yi speaker, from Fall 2009 to Summer 2010 in Boston; some data are from follow-up work with this speaker in Beijing and Boston from Summer 2010 till 2012. Data from the literature will be duly noted.

- (4) a. *mu tshɿ/a daɿ sɔ ma* horse this/that three Cl
 ‘these/those three horses.’
 c. *mu sɔ ma su* horse three Cl SU
 ‘the three horses’
 b. **si-hni tshɿ/a daɿ* horse this/that (Yi)
 d. **si-hni su* girl SU
- (5) a. *zhe/na san pi ma* this/that three Cl horse
 ‘these/those three horses.’
 b. *zhe/na nvhai-er* this/that girl (Mandarin)
 ‘this/that girl’

The contrast between (4) and (5) leads to the question of whether the alternation is a relatively ‘local’ form of syntactic variation or whether it reveals something deeper about the difference between demonstratives and article-like determiners. We will argue that the latter is the case.

The first puzzle immediately raises a second one, also to be discussed in section 2. Bare nouns in Yi behave like those in Mandarin. Not only can they freely appear in argument position, but also among their most prominent interpretations, we find the definite one and the kind oriented one. Two examples are given in (6a) and (6b).

- (6) a. *ko-lo gi o.* dinosaur extinct SFP (Yi)
 ‘The dinosaur is extinct.’
 b. *si-hni ma sini sse-vo ma i go nyi, si-hni dzi ndza.*
 girl CL and boy CL house sit, girl very beautiful
 ‘A girl and a boy are sitting in the house, the girl is very pretty.’

What this shows is that the interpretation of bare nouns in Yi includes the most central functions of the definite article. This may make one wonder as to why an overt definite article should develop in a system where its functions can be performed via other (ostensibly simpler) constructions.

Section 3 examines the behavior of Numeral Constructions (NCs for short; these are of the form *two boys* or *liang ge ren* ‘two Cl man’) at the clausal level. The reason for doing so is to show that their syntactic distribution and semantic interpretations display a formidable uniformity across languages, regardless of whether a language is a number marking one or a classifier one and whether it has an overt D. This uniform behavior of NCs will turn out to provide us with a good starting point to understand nominal arguments in Mandarin and Yi. Based on a background assumption that the function of numerals is constant cross-linguistically, we develop a general lexical approach to

numerals that provides a basis for a unified approach to NCs for both NMLs and CILs like Yi and Mandarin.

Section 4, which is central to the analytical contribution of this paper, analyzes the internal structure of NCs. It will be shown that NCs in CILs are rather different from those in NMLs in the internal domain, even though their behaviors at the clause level are so similar. We consider two analyses of the internal structure of NCs in CILs which differ in whether bare nouns denote properties or kinds. We settle on an analysis in which bare nouns in CILs are kind-referring and classifiers relate a kind-referring noun and a numeral. If Ds are universally property taking (of, say, type $\langle et, e \rangle$), we then should not expect them to combine with bare nouns in CILs. Nevertheless, we will show if some higher nominal projections above the bare noun are of a predicative type ($\langle e, t \rangle$), lexical determiners become relevant and may well develop to turn them into arguments. Language variation between NMLs and CILs in the nominal domain, as we will argue, is due primarily to two interrelated factors: what nouns denote and what low functional heads (number morphology *vs.* classifiers) denote. We further show that the proposed analysis derives the similarities and differences between Mandarin and Yi in a parametric manner. In particular, we will argue that it not only becomes unnecessary to assume a covert functional category D in Mandarin, but that it is in fact impossible to maintain such an analysis. Our analysis of Mandarin nominals will be shown to be simpler and more elegant since it avoids stipulating the presence of invisible projections that otherwise have no overt manifestation in this language. Our analysis, therefore, provides further evidence for the view which allows variation and parameterization in the obligatory status of D.

In Section 5, we lay out consequences that follow from our analysis of nominal arguments concerning the marking of definiteness via bare nouns and demonstratives in CILs (e.g. (4), (5), (6a)). We further discuss other ways of expressing definiteness in CILs. As we will see, although CILs with ‘real’ determiners like Yi are rare, there are several CILs where definiteness is somehow overtly marked. For instance, one particularly prominent case is Bangla, an Indo-European language. This language has developed a classifier system and moreover systematically marks definiteness by numeral-less classifier-noun sequences (a.k.a *bare* CLPs in the literature), with inverted word order with respect to the basic one (Bhattacharya 1999a, b; Dayal 2010, 2011b, to appear). Interestingly, in Bangla, bare nouns generally lack a definite interpretation (e.g. Dayal 2010, 2011b, to appear). What this suggests is the following generalization: definiteness marking of the type found in Bangla is seemingly incompatible with the definite use of the bare noun, while the presence of the overt definite article in Yi seems to be compatible with a definite interpretation of the bare noun. While we may not have a final account of this phenomenon, we shall see that studying Yi may pave the way to understanding what is at the basis of this alternation. In the end of Section 5, we discuss

the predictions that our analysis makes about other types of CILs and point out directions for typological investigation in the future. Section 6 summarizes and concludes.

2 Nominal arguments in CILs: similarities and differences between Mandarin and Yi

This section is dedicated to a comparison between Mandarin and Yi in the nominal domain. Their similarities and differences will be addressed in turn.

2.1 Similarities between Mandarin and Yi

First, Mandarin and Yi are both CILs: without a classifier, a numeral cannot combine directly with a noun in either languages (7a) and (8a). Mandarin NCs consist of a numeral (Num), a classifier and a noun in the order [Num-Cl-N] (7b); Yi NCs also consist of a Num, a classifier and a noun but in a different order [N-Num-Cl] (8b).²

- | | | | |
|-----|--|---|------------|
| (7) | a. * <i>san</i> <i>xuesheng</i>
three student | b. <i>san</i> <i>ge</i> <i>xuesheng</i>
three Cl student
‘three students’ | (Mandarin) |
| (8) | a. * <i>mu</i> <i>so</i>
horse three | b. <i>mu</i> <i>so</i> <i>ma</i>
horse three Cl
‘three horses’ | (Yi) |

The nouns in (7) and (8), i.e. *xuesheng* ‘student’ and *mu* ‘horse’, are conceptually count. The two classifiers *ge* and *ma* do not create a unit of measure for the nouns; instead, they simply name the units in which the entities denoted by the nouns naturally occur (e.g. see Cheng and Sybesma 1999: 515). We refer to the classifiers which do not create a unit for the nouns but simply name the standard unit coming with the nouns as ‘individual classifiers’. Besides individual classifiers, there are four more main types of classifiers: (i) container classifiers (eg. ‘bowl’), (ii) standard measure classifiers (e.g. ‘pound’), (iii) group classifiers (e.g. ‘group/team’), and (iv) partitive classifiers (e.g. ‘layer’).³

The second similarity between Mandarin and Yi concerns their NCs. They share three properties, which as we will see are stable properties of NCs across languages. These three properties are given in (9).

² There are some exceptions to this generalization. For example, in both languages, some disyllabic cardinals like ‘fifty’ (*gu-tshi* in Yi and *wu-shi* in Mandarin) can directly combine with a noun without the presence of a classifier. This phenomenon in Mandarin can be regarded as a residue from Old Chinese. According to our Yi informant, the phenomenon in Yi might be a residue from Old Yi as well. We will leave for future research the question of why certain disyllabic cardinals can allow optional classifiers.

³ Notice that the last four types of classifiers are available across languages; only the first type of classifiers—the individual classifiers are unique to CILs. See Chao (1968) for details on all types of classifiers in Mandarin.

- (9) Three properties of NCs in Mandarin and Yi
- (i) Property 1: They have predicative use
 - (ii) Property 2: They can be used as restrictors of demonstratives and quantifiers.
 - (iii) Property 3: They are construed as indefinites in argument position

We will elaborate on them in turn. The first property is exemplified in (10).

- (10) a. *Zhangsan, Lisi, Wangwu shi san ge xuesheng.* (Mandarin)
 Zhangsan, Lisi, Wangwu be three Cl student
 ‘Zhangsan, Lisi and Wangwu are three students.’
- b. *Aka sini Ako sse-vo nyip ma nyu.* (Yi)
 Aka and Ako boy two Cl be
 ‘Aka and Ako are two boys.’

The predicative use of NCs (in which they are presumably construed as being of a predicative semantic type, say <e,t>) is also probably at the basis of their capacity to serve as the restrictors of demonstratives (Dem) and quantifiers, as exemplified in (11).

- (11) a. *zhe/na wu ge nv-sheng a'. mei liang ge xuesheng* (Mandarin)
 this/that five Cl girl every/each two Cl student
 ‘those five girls’ ‘every two students’
- b. *mu tshɿ/a daɿ so ma b'. ssox sse cyp ma dzi³ a² dzi* (Yi)
 horse this/that three Cl student one Cl every
 ‘these/those three horses.’ ‘every student’

In addition to their predicative use, Mandarin and Yi NCs can also appear in argument position with an indefinite interpretation. They receive a narrow scope existential reading in (12a) and (12b) and can appear in existential sentences, as in (12a') and (12b').

- (12) a. *wo mai le san zhang zhuzi* (Mandarin)
 I buy Asp three Cl desk
 ‘I bought three desks.’
- a'. *fangjian li you liang ge ren*
 room inside have two Cl man
 ‘There are two men in the room.’
- b. *nga zhuop-zyr so ma vy lo.* (Yi)
 I desk three Cl bought
 ‘I bought three desks.’

Example (12) (Continued)

b'. mu sɔ ma a-di dzu
Horse three Cl there have
'There are three horses'

The third main similarity between Mandarin and Yi is that they both allow bare arguments. In particular, Mandarin bare nouns receive a kind interpretation when occurring with a kind level predicate (e.g. Krifka 1995) (13i); they receive a generic reading in generic sentences (e.g. Cheng and Sybesma 1999; Yang 2001) (13ii); in episodic sentences, they can receive either a definite or an existential reading in both subject and object position (e.g. Yang 2001) (13iii).

(13) Mandarin Bare nouns

(i) bare nouns with kind level predicates

xiong jue-zhong le.
bear vanish-king Asp
'The bear is extinct.'
(Krifka 1995, ex. 1a)

(ii) bare nouns in generic sentences

gou ai jiling.
dog very smart
a. 'Dogs are intelligent.'
b. 'The dog(s) is/are intelligent.'
(Yang 2001, ex. 12c)

(iii) bare nouns in episodic sentences

a. waimian gou zai jiao.
outside dog Prog bark
i. 'Outside, dogs are barking.'
ii. 'Outside, the dog(s) are/is barking.'
(Yang 2001, ex. 31)
b. fangjian li zuo zhe yi ge nansheng yi ge nvsheng,
room inside sit Prog one Cl boy one Cl girl
nansheng kan-qi-lai hen nianqing.
boy look very young
'There is a boy and a girl sitting in the room; the boy looks very young.'

It is worth emphasizing that, as discussed in Yang (2001) and X. Li (2011), the scope properties of bare nouns in Mandarin (and as we shall see shortly in Yi) regarding their indefinite reading resemble those of English bare plurals (e.g. dogs) rather than those of lexical indefinites (e.g. a/some dog(s)). Since Carlson (1977b), it has been known that English bare plurals systematically differ from regular indefinites in their scope properties. Bare plurals have *only* a narrow scope existential reading, while regular indefinites can receive both a wide and a narrow scope reading under negation and under

other scope bearing elements. For example, the bare plural ‘dogs’ in (14a) only receives the reading where it takes scope under negation, whereas the regular indefinite ‘a dog’ in (14b), can have either a narrow scope or a wide scope reading.

- (14) a. I didn’t see dogs. $\neg > \exists / * \exists > \neg$
 b. I didn’t see a dog. $\neg > \exists / \exists > \neg$

In addition, indefinites can escape all islands for extraction (Ruys 1992: 102-103), as exemplified in (15a). Unlike definites which always take wide(st) scope, indefinites can escape islands *without* having the widest sentential scope (Farkas 1981; Ruys 1992; Abusch 1994). An example to demonstrate this point is given in (15b). The indefinite ‘a student of his’ can escape the adjunct island and have a scope narrower than the sentence initial scope-bearing element ‘every’. The scope of this type (e.g. (15bi)) is usually referred to as *Intermediate Scope*.

- (15) a. If John hires two workers, he will be in good shape.
 i). Wide scope: two workers > if ii). Narrow scope: if > two workers
 b. Every professor will rejoice [if a student of his cheats on the exam].
 i). Intermediate Scope: every > a student of his > if
 ii). Narrow Scope: every > if > a student of his (Ruys 1992)

In Mandarin, if we set aside the definite interpretation, bare nouns behave just like English bare nominals: they can receive a narrowest existential interpretation (16iia) and (17iia), but not a wide scope reading (16iib) nor an intermediate scope reading (17iib) (e.g. see Yang 2001, X. Li 2011).

- (16) *mei ge ren dou zai kan xin.*
 every Cl man all Prog read letter
 i. ‘Everybody is reading the letter.’ [definite]
 ii. ‘Everybody is reading letters.’
 a. Narrow scope: [$\forall >$ some letters] b. *Wide scope: *[some letters/a letter > \forall]
 (X. Li 2011, ex 19b with slight modification)

- (17) *dabufen yuyanxuejia dou kan-guo mei-ge [jiejue wenti de] fenxifangfa.*
 most linguist all look-Asp every-CL solve problem DE analysis
 i. ‘Most linguists have looked at every analysis that solves the problem.’
 ii. ‘Most linguists have looked at every analysis that solves problems.’
 a. Narrow scope: [most > $\forall >$ problems]
 b. *Intermediate scope: *[most > some problems/a problem > \forall]
 (Yang 2001, ex. 26)

Turing to Yi, we note that its bare nouns behave just like those in Mandarin with respect to semantic interpretations and syntactic distributions, as exemplified in (18).

(18) Yi bare nouns

(i) bare nouns with kind-level predicates

ko-lo gi o.
 dinosaur extinct SFP
 ‘Dinosaurs are extinct.’

(ii) bare nouns in generic sentences

Aka khu tee si ko-ba.
 Aka dog scared very much
 ‘Aka is scared of dogs very much.’

(iii) bare nouns in episodic sentences

- a. *Aka thu-zi vi bo o.*
 Aka book buy go SFP
 i. ‘Aka went to buy books.’
 ii. ‘Aka went to buy the book(s).’
- b. *ne si-hni shyp ngop ddu ggep la go-li, nga khat ox.*
 you girl bring my home play come if, I happy SFP
 i. ‘If you bring the girls to hang out in my house, I will be happy.’
 ii ‘If you bring girls to hang out in my house, I will be happy.’
 (a). Narrow scope existential [if > girls] (b). *Wide scope [*some girls > if]
- c. *si-hni ma sini sse-vo ma i go nyi, si-hni dzi ndza.*
 girl CL and boy CL house sit, girl very beautiful
 ‘A girl and a boy are sitting in the house, the girl is very pretty.’

Summarizing up so far, Yi and Mandarin behave rather similarly in the nominal domain, i.e. a classifier is obligatory when numerals combine with nouns, their NCs share three common properties, and their bare nouns behave alike regarding the syntactic distribution and semantic interpretations; however, we do observe two dimensions of variation that distinguish them from each other.

2.2 Differences between Mandarin and Yi

First, Mandarin and Yi differ in whether numeral-less classifier phrases (a.k.a *bare CLPs* in the literature), are freely allowed. This is a dimension along which CLs vary significantly (see Cheng and Sybesma 1999, 2005; Simpson 2005). For instance, Cantonese freely allows bare CLPs in both subject and object position with a *definite* interpretation (e.g. Cheng and Sybesma 1999) (19), Mandarin only allows *indefinite* ones in a restricted way, i.e. in a position immediately following a verbal element or a stressed nominal element (cf. e.g. Lü 1944) (20), and Korean/Japanese simply bans them (21).

- (19) a. *bzek gau zungji sek juk.* (Cantonese)
 CL dog like eat meat
 ‘The dog likes to eat meat.’
 b. *Ngo zungji tong zek gau waan.*
 I like with CL dog play
 ‘I like to play with the dog.’ (Cheng and Sybesma 1999: 511 (6))
- (20) a. **ge xuesheng yinggai haohao xuexi.* (Mandarin)
 Cl student should good-good study
 Intended: ‘A student should study hard.’
 b. *wo jiao le ge nongmin, *ge xuesheng he ge gongren.*
 I teach Asp Cl farmer, Cl student and Cl worker
 Intended: ‘I taught a farmer, a student and a worker.’
- (21) a. **soi mali-ka swuley-lul kkul-ko iss-ta.* (Korean)
 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 a student.’

Compared with bare CIPs in other CILs, Yi bare CIPs behave quite differently in that although they are freely allowed in both subject and object position, they *only* receive an *indefinite* interpretation (22).

- (22) a. *tsho ma dza dzu ndzo.* [indefinite]
 person CL rice eat Progressive
 ‘A person is having meal.’
 b. *tshi mu ma su bo o.* [indefinite]
 3s horse Cl look-for go SFP
 ‘He went to look for a horse.’
 c. *si-hni ma sini sse-vo ma i go nyi, #si-hni ma dzi ndza.* [*definite]
 girl Cl and boy Cl house sit, girl Cl very beautiful
 Intended: ‘A girl and a boy are sitting in the room; the girl is very pretty.’
 d. *a-nyie ma a-hie yo yie.* [generic]
 cat CL mouse catch should
 ‘A cat should catch mice.’
 e. **ko-lo ma gi o.* [*kind]
 dinosaur CL extinct SFP
 Intended: ‘Dinosaurs are extinct.’

These examples illustrate how Yi bare CIPs can freely appear in subject and object position in episodic sentences, receiving an indefinite interpretation (22a, b); this is also supported by the fact that they cannot be used anaphorically (22c). They can appear in generic sentences with a generic interpretation (22d) but cannot occur with kind level predicates (22e). In other words, they seem to behave like singular indefinites (like *a cat*) in English.

The second dimension of variation between Mandarin and Yi lies in the existence of an overt definite determiner. Like the majority of CILs, Mandarin lacks overt determiners like ‘the’ or ‘a’ although it does have, of course, demonstratives (e.g. (5)). Yi, in contrast, has a morpheme *su* that has all the functions that definite determiners like English ‘the’ have (e.g. Jiang and Hu 2010; Jiang 2012). *Su* appears in the final nominal position; it turns indefinite NCs and indefinite bare CIPs (23a, b) into definites (23a', b').

- | | | |
|------|---|--|
| (23) | a. <u><i>zhuop-zyr sɔ ma</i></u>
desk three CI
‘three desks’
b. <u><i>tsho ma</i></u>
person CL
‘a person’ | a'. <u><i>zhuop-zyr sɔ ma su</i></u> (Yi)
desk three CI SU
‘the three desks’
b'. <u><i>tsho ma su</i></u>
person CL SU
‘the person’ |
|------|---|--|

It is important to point out that the post-nominal morpheme *su* differs from demonstratives in Yi, which appear in the position between nouns and classifiers, and cannot co-occur with them (24).

- | | | |
|------|--|------|
| (24) | a. <i>zhuop-zyr tshɿ/a daɿ sɔ ma (*su)</i>
desk this/that three CI SU
‘these/those three desks’
b. <i>zhuop-zyr tshɿ/a daɿ ma (*su)</i>
desk this/that CI SU
‘this/that desk’ | (Yi) |
|------|--|------|

In addition to being used referentially, *su*-phrases can also be used anaphorically to pick the maximal entity in an antecedent clause (25), much like definites in English. The anaphoric use of *su*-phrases is identical to that of bare nouns (18iic). What this shows is that definiteness can be expressed either via bare nouns or *su*-phrases in this language.

- (25) a. *si-hni nyip ma sini sse-vo ma i go nyi, si-hni nyip ma su dzi ndza.*
 girl two Cl and boy Cl house sit, girl two Cl SU very beautiful
 ‘Two girls and a boy are sitting in the room; the two girls are very pretty.’
 b. *si-hni ma sini sse-vo ma i go nyi, si-hni ma su dzi ndza.*
 girl Cl and boy Cl house sit, girl Cl SU very beautiful
 ‘A girl and a boy are sitting in the room; the girl is very pretty.’

Furthermore, *su*-marked NCs can occur in generic sentences as well; interestingly, they can be interpreted in two ways. For example, in (26), *mu sɔ ma su* ‘horse three Cl su’ can be interpreted as definites ‘the three horses’ (26i) or generics ‘any three horses’ (26ii).

- (26) *mu sɔ ma su sɔ tu va bi.*
 horse three Cl Su three thousand dollar give
 i. ‘The three horses cost three thousand dollars.’ [definite]
 ii. ‘(Any) Three horses cost three thousand dollars.’ [generic]

One may wonder whether *su* could be a case marker. We only saw examples with *su*-phrases in argument position, so it could be the case that *su* is a nominative or an accusative marker. However, it turns out that *su*-phrases can freely co-occur with overt oblique case markers (27); this excludes the possibility that *su* is a structural case marker.

- (27) *Aka bbap-ga ma su ta la.*
 Aka village Cl SU Ablative come
 ‘Aka comes from the village.’

The data presented above show that *su* contributes to the definite interpretation of Yi NCs and bare ClPs. *Su*-phrases can be used not only referentially but also anaphorically to refer to the maximal entity in the antecedent clause. The maximality effect that *su*-phrases exhibit is a manifestation of their definite nature. We also saw that phrases containing *su* have generic uses in generic sentences. Structurally, *su* appears in the post-nominal position, differing from demonstratives in Yi that occur in the position between nouns and classifiers. Based on such data, it is legitimate to claim that *su* has the same role as elements that are uncontroversially assumed to be definite determiners (e.g. English ‘the’) and that it should be analyzed as such.

Interestingly, although *su* combines freely with NCs and bare ClPs, it cannot combine directly with bare nouns (and demonstratives cannot either), as illustrated in (28). In order to make (28) grammatical, a classifier needs to be added after the noun, or *su* has to be removed.

- (28) **si-hni su dzi ndza*. (Yi)
 girl SU very beautiful

To summarize this section, we saw three similarities and two dimensions of variation between Mandarin and Yi with regard to their nominal arguments. It was shown that their similarities are quite uniform across CILs. CILs differ a great deal along the first dimension of variation, while variation along the second dimension -- the existence of an overt determiner -- is rarer (Yi appears to be the only CIL with an overt article-like determiner found so far). In the next section, we will examine the behavior of NCs at the clausal level; this will turn out to provide us with a good starting point to understand nominal arguments in Mandarin and Yi.

3 The external syntax and semantics of NCs: A baseline proposal

This section examines the external syntax and semantics of NCs like *three cats/san zhi mao* ‘three Cl cat’. The reason for doing so is to show that their syntactic distribution and semantic interpretations display a formidable uniformity across languages, regardless of whether the language is a number marking one or a classifier one and whether or not it has an overt D. Presumably, this behavior is universal, hence the result of combining a numeral with a nominal should receive a uniform analysis across languages. We will sketch such analysis, and this will constitute our baseline. From it, we hope to arrive at understanding the inner components of nominals by ‘subtraction’, so to speak.

3.1 Some tendentially universal patterns of NCs

In Section 2, we saw three properties of NCs in Mandarin and Yi (cf (9)); these properties are also attested in NMLs regardless of ‘how much D’ they have. Here they are illustrated for English, French and Russian:

- (29) English NCs
- (i) Property 1: Predicative uses
 John and Bill are two boys.
 - (ii) Property 2: Restrictors of definites/quantifiers
 - a. John hired the/those two boys.
 - b. I made different assignments for every four students.
 - (iii) Property 3: Argumental indefinites and existential readings
 - a. John hired two workers.
 - b. There are two workers.

- (30) French NCs:
- (i) Property 1: Predicative uses
Ils sont deux étudiants
 They are two students
 ‘They are two students.’
 - (ii) Property 2: Restrictors of definites/quantifiers
Les /ces deux chats
 The/these two cats
 ‘The/these two cats’
 - (iii) Property 3: Argumental (indefinite): narrow scope existential reading

a. <i>J’ai vu <u>deux chiens</u>.</i>	b. <i>Il y a <u>deux chiens</u>.</i>
I saw two dogs	there have two dogs
‘I saw two dogs’	‘There are two dogs’

- (31) Russian NCs:
- (i) Property 1: Predicative
Ivan i Anna dva studenta
 Ivan and Anna two students
 ‘Ivan and Anna are two students.’
 - (ii) Property 2: Restrictors of definites/quantifiers
Eti pjat’ koshek
 these five cats
 ‘these five cats’
 - (iii) Property 3: Argumental (indefinite): narrow scope existential reading

a. <i>V komnate <u>pjat’ sobak</u>.</i>	b. <i>Ja videla <u>pjat’ sobak</u>.</i>
in room five dogs	I-nom see.past five dogs
‘There are five dogs in the room.’	‘I saw five dogs’

Now, English, French and Russian differ greatly in DP-structure. English has overt Ds but also allows bare arguments. French has overt Ds and systematically disallows bare arguments. Russian has no overt Ds (i.e. articles). Yet, their NCs appear to be identical with respects to the properties in (29)-(31), shared also by CILs, as we saw. This identity extends also to further properties of NCs, namely:

- (32) Three additional common properties of NCs across languages
- (i) Property 4: Lack of anaphoric uses
 - (ii) Property 5: Long-distance scope interpretation and island-escaping ability
 - (iii) Property 6: Generic construals

We illustrate schematically these properties for English in (33)-(35). Parallel facts for French, Russian, Mandarin and Yi are provided in the Appendix.

- (33) Lack of anaphoric uses:
- a. John bought two dogs and five cats. #*(The/Those) five cats are very expensive.*
- (34) Long distance scope taking of NCs
- a. [If John hires two workers] he will be in good shape.
 - i). **Wide Scope:** [two workers > if]
 ‘There are two workers such that if John hires them, he will be in good shape.’
 - ii). **Narrow Scope:** [If >two workers]
 ‘If Joohn hires any two workers, he will be in good shape.’
 - b. Every professor will rejoice [if three students of his do well on the exam].
 - i). **Intermediate Scope:** [Every >three students of his > if]
 ‘For every professor x there are three students such that x will rejoice if those three students do well.’
 - ii). **Narrow Scope:** [Every > if >three students of his]
 ‘Every professor will rejoice if any group of at least three students will do well.’
- (35) Generic uses of NCs in English
- a. Three boys can lift the piano.
 - b. Two canaries can be kept in the same cage if it is large enough.
 - c. Six apples cost one dollar. (Krifka et al 1995, ex. 62, 99)
 - d. Episodic counterparts of (a)-(b):
 - i. Three boys are lifting a piano.
 - ii. Two canaries were kept in the same cage yesterday.

The sentences in (35a-c) all report general properties of the NCs; they all constitute generic sentences in the sense of Krifka et al (1995: 2-3). For instance, ‘three boys’ in (35a) does not refer to three specific boys or to a *particular* event. This sentence expresses rather a quantificational statement *over events* and reports a general property that any three boys claimed to have. They contrast with the episodic sentences in (35d).

So, NCs appear to have identical syntactic distribution and semantic interpretation with respect to all of the six properties considered in the above data set, regardless of whether a language (obligatorily) requires overt Ds to argumentize nouns and whether a language allows numerals to combine directly with nouns. These facts lead us to the following generalizations on NCs.

- (36) Tendentially universal properties of NCs:
- i. NCs are always both predicative and argumental.
 - ii. In their argumental role, they are always indefinites (with a peculiar scope).
 - iii. In their non-argumental role, they can combine with a definite element/marker and then (and only then) they become definite.

Example (36) (Continued)

iv. Properties i-iii are stable, regardless of whether or not a language has (overt) Ds and whether or not a language is a NML or a CIL.

The question is then how would analysis of these very general facts look like.

3.2 The semantics and the syntax of numerals: a general approach to NCs

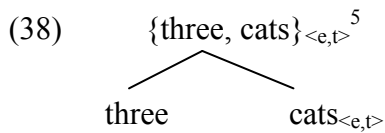
Our goal of this section is to outline a general approach to NCs that can capture the uniform behavior of NCs across languages. This approach should provide a basis for a unified approach to NCs for both NMLs and CILs, like Yi and Mandarin.

A reasonable background assumption is that the function of numerals is constant cross-linguistically. As is widely known, numerals can be analyzed as being of three basic types. First, numerals can denote properties of type $\langle e, t \rangle$ (37i); second, numerals can be property modifiers, type $\langle \langle e, t \rangle, \langle e, t \rangle \rangle$ (37ii); third, numerals can denote entities, type $\langle e \rangle$ (37iii); perhaps one of these types is basic, while others are derived.

(37) Basic types of numerals

- i. properties: *Those cats are* $\underline{3}_{\langle e, t \rangle}$.
- ii. property modifiers: *those* $\underline{3}_{\langle \langle e, t \rangle, \langle e, t \rangle \rangle}$ *boys*.
- iii. entities: a. $\underline{5}_{\langle e \rangle}$ *is my lucky number*. b. *John is* $\underline{5}_{\langle e \rangle}$ *feet tall*.⁴

While it is controversial how the three uses are related to each other and which use of numerals above is more basic, what matters to our current study is the type of NCs should be $\langle e, t \rangle$. The reason is straightforward: NCs can always be predicates qua restrictors of definites and quantifiers (as we saw in Section 3.1). In addition, when numerals can combine directly with a noun, what everyone would agree is that the noun should be property-denoting, type $\langle e, t \rangle$, given that numerals require the nominal that they combine with to denote a set containing atoms/groups. Hence, a rough structure of NCs would be the one below.



⁴ Numerals have been analyzed as ranging over degrees (cf. e.g. Kennedy 1999; Kennedy and McNally 2005), so that, (37iiib) can be analyzed as something like ‘the degree of John’s height measured in feet is *five*’.

⁵ In the structure in (38), there could be a functional head like the singular-plural head that merges with the bare noun first, forming $\{ -s, \text{cat} \}$; then $\{ -s, \text{cat} \}$ further merges with the numeral, resulting $\{ \text{three}, \{ -s, \text{cat} \} \}$. We will discuss that in the end of this section.

In (38), we represent the structure of ‘three cats’ in the framework of *bare phrase structure* as in Chomsky (1994 et seq). According to this framework, the basic structure-building operation that comes ‘free’ (in that it is required in some form for any recursive system) is *Merge*. Merge can be viewed as a function that takes two elements, say α and β , and creates a new one consisting of the two; the simplest element constructed from α and β is the unordered set $\{\alpha, \beta\}$ (see Chomsky 2004: 108).⁶ For instance, in (38) Merge operates on the lexical items ‘three’ and ‘cats’ and maps them into a set $\{\text{three}, \text{cats}\}$.⁷

Regarding the semantics of numerals, we will simply adopt the analysis of numerals as privative (non-subjective) adjectival modifiers of type $\langle \text{et}, \text{et} \rangle$ as proposed in Ionin and Matushansky 2006. We illustrate this analysis of numerals in (39).⁸

(39) Semantics of numerals: as modifiers (*first version*)

Numerals = $\lambda P \lambda x [n(P)(x)]^9$ $\langle \text{et}, \text{et} \rangle$ (Ionin and Matushansky 2006)

We consider two main reasons for adopting this analysis of numerals in (39). First, this analysis is the simplest one in terms of types (e.g. in any intro to semantic course, everyone would give the same type as $\langle \text{et}, \text{et} \rangle$ for numerals when presented with the structure in (39)). Second, the analysis can explain not only simple numerals but also complex numerals like *five hundred* (e.g. see Ionin and Matushansky 2006 for details). Note that, even if the analysis in (39) is shown to be inadequate or incorrect, other analyses of numerals are still compatible with what we are going to propose.¹⁰

The semantics of numerals in (39) naturally captures two of the cross-linguistic properties of NCs: the predicative use and the use as restrictors of quantifiers and definites. We illustrate them in (40) and (41) with examples from English.

⁶ When *Merge* maps two elements α and β into a set of the two $\{\alpha, \beta\}$, some mechanism has to be put in place to determine which of the two elements should be the head of the phrase. Perhaps this can be done by the type theory or some feature-based theory; this is a general concern, and the reader should feel free to adopt his/her favorite analysis.

⁷ Although we have adopted the framework of *bare phrase structure* (Chomsky 1994, 1995 et seq) to illustrate the structure of NCs, not much would change if we adopted the standard X-bar theory.

⁸ The analysis of NCs that we are going to propose is also compatible with the view that treats numerals as restrictive adjectival modifiers (as in Link 1983; Bale et al 2010, among others)

⁹ The semantics of the numeral n could be simply analyzed as the one below:

(1) $[[n \text{ men}]] = n(\text{men}) = \lambda P \lambda x \exists Q [Q \subseteq \text{AT}(P) \wedge |Q| = n \wedge x = \cup Q](\text{men})$

The analysis above is not the same as the one proposed in Ionin and Matushansky 2006; however, it works for our purpose of discussion and allows us to capture complex numerals like *three hundred men* via recursive syntax and compositional semantics. We refer the readers to Ionin and Matushansky 2006 for a more explicit semantics of the numeral n in the same spirit.

¹⁰ As we saw earlier, there are several possibilities for the semantics of numerals, but we will not consider all of them because the details are not essential for the purpose of the current study and the differences among these possibilities do not make much difference for us.

- (40) a. John, Bill and Tom are *three students*. (as *predicates*)
 b. $[[\text{three}]] = \lambda P [\text{three} \langle \langle e, t \rangle, \langle e, t \rangle \rangle (P)]$ $\langle \langle e, t \rangle, \langle e, t \rangle \rangle$
 c. $[[\text{boys}]] = \text{boys}$ $\langle e, t \rangle$
 d. $[[\text{three boys}]] = \text{three (boy)}$ $\langle e, t \rangle$
 e. $[[\text{John, Tom, and Bill are three boys}]] = \text{three (boy) (John} \cup \text{Bill} \cup \text{Tom)}$
- (41) a. *The/Every three boys* brought water. (as *restrictors* of definites/quantifiers)
 b. $[[\text{three boys}]] = \text{three (boy)}$ $\langle e, t \rangle$
 c. $[[\text{the three boys}]] = \iota \text{ three (boy)}$ $\langle e \rangle$
 d. $[[\text{the three boys brought water}]] = \text{brought} (\iota \text{ three (boy), water})^{11}$

In addition to their predicative use, NCs across languages can freely appear in argument position as we saw in Section 3.1. In their argumental role, NCs are always indefinites, with either a narrow scope existential reading or a long-distance scope reading. Concerning the existential force of NCs, the traditional way is to derive it via a generalized quantifier (GQ) variant of NCs. This quantificational force of numerals is widely assumed to be related to their modifier use: it can be obtained via a global existential closure (Heim 1982), a covert existential quantifier (Link 1983, 1987; Krifka 1999), or a type-shifting principle (Partee 1986; Landman 2003). We remain agnostic as to how exactly this quantificational force of numerals is gained. Below we demonstrate one of the possibilities of deriving the quantificational force of NCs from their modifier use.¹²

- (42) GQ use of NCs:
 Numerals = $\lambda P \lambda Q \exists x [n (P) (x) \wedge Q (x)]$ $\langle \langle e, t \rangle, t \rangle$
 i. $[[\text{three boys}]] = [[\text{three}]] ([[\text{boys}]]) = \lambda P \lambda Q \exists x [\text{three} (P) (x) \wedge Q (x)] (\text{boy}) = \lambda Q \exists x [\text{three}(\text{boy})(x) \wedge Q(x)]$
 ii. $[[\text{three boys left}]] = [[\text{three boys}]] ([[\text{left}]]) = \lambda Q \exists x [\text{three}(\text{boy})(x) \wedge Q (x)] (\text{left}) = \exists x [\text{three}(\text{boy}) (x) \wedge (\text{left}) (x)]$

Concerning the long-distance scope interpretation of NCs, the GQ analysis of NCs encounters scope problems since it only derives a *narrow scope existential* interpretation. The peculiar scope behavior of NCs requires an analysis that is distinguishable from standard quantification, and an analysis of indefinites in terms of choice functions has

¹¹ It is worth noting that a phrase like *every three houses* does receive a special interpretation; in particular, the quantifier must range over non-overlapping triplets. However, we cannot get into the details in the current study; for the purpose of our discussion, it is suffice to point out that NCs must have a type compatible with the quantifier.

¹² According to the Principle of Same-Type Coordination in Partee 1987, only categories of the same semantic type can be coordinated. Thus when coordinated with ‘every girl’ in *I met three boys and every girl*, the NC ‘three boys’ must also be regarded as a generalized quantifier, type $\langle \langle e, t \rangle, t \rangle$.

been pursued in Reinhart (1997), Winter (1997), and Kratzer (1998). In these works, indefinites are assumed to involve a function variable in their semantics that assigns an individual to the restriction of the predicate (c.f. Winter 1997: 409); this function is a choice function which is subject to existential closure. The existential quantification associated with the choice function is assumed to apply at any scope site (see Winter 1997: 409-411 for details). Take (43) for instance, if the existential closure is performed within the adjunct clause (43ia), the sentence receives a narrow scope interpretation and can be paraphrased as (43ib). If the existential closure is applied outside the conditional (43iia), a wide scope interpretation is gained; the sentence can be paraphrased as (43iib), where $CH(f)$ means that f is a choice function.

- (43) [If **one woman** comes to the party] John will be glad.
- i). Narrow scope interpretation
 - a. $[\exists f [CH(f) \wedge \text{come}(f(\text{woman}))]] \rightarrow \text{glad}(\text{John})$
 - b. John will be glad if there is *any* possibility to pick a woman who comes to the party.
 - ii). Wide scope interpretation
 - a. $\exists f [CH(f) \wedge [\text{come}(f(\text{woman})) \rightarrow \text{glad}(\text{John})]]$
 - b. There is a choice function such that John will be glad if the *women it picks* comes to the party. (Winter 1997: 411, ex (16, 16') and (17, 17'))

Given that the interpretation of the indefinite is determined by *the stage* at which existential closure is applied, no mechanism extracting the indefinite out of the island is involved; thus *no* island constraint is violated (c.f. Winter 1997: 411).

We adopt the choice function analysis of numeral indefinites because it not only provides a uniform account of different indefinites (*a*-, *some*-, and numeral indefinites), but also it captures the long-distance scope behavior of NC found cross-linguistically. We refer the choice function variable that is subject to existential closure as ' f_{\exists} '.

Now, a broad question immediately arises, namely, how we can relate the predicative use of NCs to its argumental use. A simple way to relate these two different uses is to attribute the choice function variable in numerals as part of their lexical semantics. In particular, each numeral n (of type $\langle\langle e, t \rangle, \langle e, t \rangle\rangle$) can be viewed as a lexically predictable alternant of type $\langle\langle e, t \rangle, e\rangle$, formed via a lexical rule, as illustrated in (44) (e.g. see also Dayal to appear for a similar view).

- (44) A lexical view of the choice function variable
- $$\text{Numeral}_{\langle\langle e, t \rangle, e\rangle} = \lambda P f_{\exists} (n(P)) \quad \langle\langle e, t \rangle, e\rangle$$

We consider one empirical argument and one theoretical argument for this lexical view of choice functions. Its empirical motivation comes from the stable cross-linguistic property

of NCs. As we saw in Section 3.1, although languages may differ greatly with respect to argumentizing bare nouns (e.g. (obligatorily) via an overt D or not via an overt D at all), they systematically allow NCs to freely appear in argument position with a peculiar scope interpretation. The stable argumental behavior of NCs across languages suggests that NCs must have a different source to form arguments from bare nouns which vary cross-linguistically. Such a source that enables all NCs to behave in a rigid way, if it is not from the narrow syntax, is very likely to lie in the lexical entry of numerals. Turning to the theoretical consideration of this lexical view, it provides a unified approach to NCs to both NMLs and CILs. The choice function variable built into the lexical entry of numerals is not subject to the existence of any functional head which might be subject to parameterization (e.g. D). In other words, the lexical view of numerals can reduce unnecessary theoretical assumptions and parameterization regarding NCs in the syntax, helping us to maintain a minimalist fashion of linguistic inquiry.

Admittedly, one can accommodate the choice function in the syntax rather than in the lexical entry of numerals. For example, the choice function can be tied to some functional head that presumably is available in all NMLs and CILs given that their NCs exhibit uniform behaviors. Or, we can assume that the choice function is linked to some functional head which might be subject to parameterization, e.g. null D (as pursued in Winter 2001, 2005 and Ionin & Matushansky 2006 for English); then we need to make other assumptions about where the choice function is realized in languages without such a functional head (i.e. languages without Ds). Even if we assume that languages without overt determiners also project null DPs in the syntax, we then need to explain why bare nouns, such as English *dogs*, do not have the same scope behavior as NCs since they might as well be argumentized via a null D. As we saw in (14a), (16) and (17), bare nouns in English, Mandarin, and Yi only receive a narrowest scope existential interpretation. Of course, one can assume two types of null Ds: one is linked to the choice function variable and merges with NCs, leading to the long-distance scope behavior; the other merges with bare nouns and only contributes to the narrowest existential reading (e.g. the Longobardi's 1994 style of null D). However, such a parametric approach of Ds not only doesn't show advantage over the lexical approach of numerals, but also it inevitably increases stipulations in the syntax.¹³ No matter what assumption about NCs one makes, it should capture their unvarying argumental behavior across languages as well as their idiosyncratic scope behavior. The proposed lexical analysis of choice function straightforwardly captures this cross-linguistic uniformity; it also displays theoretical advantages over a narrow syntax based analysis which inevitably would be

¹³ A potential question may arise here, namely, why the source of choice functions cannot be a covert semantic operation like type-shifting. This semantic assumption is theoretically plausible. However if we allow such a covert semantic operation for NCs in French, say, one would like to know why it is not available for bare nouns. Admittedly, one can always make some assumptions to explain this puzzle; however, we believe that adding further assumptions would inevitably result in a more stipulative theory than the proposed lexical analysis.

more stipulative. Consequentially, numerals, under our analysis, are lexically ambiguous, as summarized in (45).

- (45) Lexical rule of ambiguous numerals (final version)¹⁴
 Numeral_{<<e,t>, <e,t>>} = $\lambda P[n(P)]$
 If $[[\alpha]] \in \text{Num}_{<<e,t>, <e,t>>}$, then $\lambda P \exists (\alpha(P)) \in \text{Num}_{<<e,t>, e>}$
 \exists is then subject to existential closure at arbitrarily chosen scope sites.

The lexical view in (45) allows us to account for the predicative use of NCs; it also explains why NCs are always indefinite with a unique scope behavior. Consequences are demonstrated below with two examples from English.

- (46) a. John saw three boys. (indefinite)
 b. $[[\text{three}]] = \lambda P \exists [\text{three}_{<<e,t>, <e,t>>}(P)]$ <<e,t>, e>
 c. $[[\text{three boys}]] = \exists (\text{three}(\text{boy}))$ <e>
 d. $[[\text{John saw three boys}]] = \exists f [\text{CH}(f) \wedge \text{saw}(\text{John}, f(\text{three}(\text{boy})))]$
- (47) a. If John hires three boys, he will be in good shape. (long-distance scope)
 b. $[[\text{three boys}]] = \exists (\text{three}(\text{boy}))$ <e>
 c. Wide Scope Interpretation: [three > if]
 $\exists f [\text{CH}(f) \wedge [\text{hire}(\text{John}, f(\text{three}(\text{boy}))) \rightarrow \text{in good shape}(\text{John})]]$
 d. Narrow Scope Interpretation: [if > three]
 $[\exists f [\text{CH}(f) \wedge \text{hire}(\text{John}, f(\text{three}(\text{boy})))]] \rightarrow \text{in a good shape}(\text{John})]$

Now we turn to the last use of NCs, i.e. the use in generic sentences, with the aim of understanding how genericity comes about. As argued in details in Krifka et al 1995: 14-21, ‘the locus of the genericity in generic sentences is *not* in the nominal subject but rather in the sentence itself’. The authors show that generic sentences put no restriction on what types of nominal phrases may occur in them (see Krifka et al. 1995 for arguments). Some examples from English are given below.

- (48) Different types of nominals in generic sentences
- | | |
|---|---|
| a. <u>John/My brother</u> drinks Whiskey. | b. <u>Every Professor</u> drinks whiskey. |
| c. <u>The Italian</u> drinks whiskey. | d. <u>Professors</u> drink whiskey. |
| e. <u>Milk</u> is healthy. | f. <u>A boy</u> should not cry. |
- (Krifka et al. 1995: 8-55, ex (19), (55e), (62), (99) with slight modification)

¹⁴ In addition to allowing us to account for the long-distance behavior of NCs, the proposed lexical view of choice functions also allows us to derive the GQ variant of NCs (e.g. (42)), as illustrated below.

(2) The GQ use of NCs: Numerals = $\lambda P \exists f [P(f(\lambda x n(P)(x)))]$

The variety of nominals acceptable in generic sentences makes it implausible that this type of genericity is determined solely by the meaning of an NP. The generic meaning, as assumed in Lawler 1972; Dahl 1975; Carlson 1977a, b, is contributed by an operator *Gen* which quantifies over situations as well as objects. As for the semantics of this *Gen* operator, there are various views, each with its pros and cons, as reviewed and discussed in Krifka et al 1995. Generally, the *Gen* operator can be viewed as a universal quantifier (\forall) quantifying over situations/possible worlds and objects (e.g. see Kratzer 1981; Chierchia 1995); this is the view that we adopt for the semantics of generic sentences. Syntactically, the *Gen* operator needs to be located in a *structurally high enough* position to quantify over both the external argument and the internal argument. Accordingly, it should be introduced at a position above *vP* (assuming the VP-internal Subject Hypothesis; see, e.g., Koopman and Sportiche 1991). Here we will simply assume that the *Gen* operator merges with IP since the detail does not make much difference for the purposes of the current discussion (or see Chierchia 1995 for the assumption that *Gen* merges with *AspP*).

We have shown how genericity is understood semantically and structurally above; it is, nevertheless, important to point out that there still remain lots of poorly understood issues about it. For example, in some cases NCs are unacceptable in generic sentences (49b); while in other cases, they are well-accepted (50).¹⁵

- (49) a. Cats are beautiful when they have white fur.
 b. ?Twelve cats are beautiful when they have white fur. (Krifka et al 1995: 35)
- (50) a. Two canaries can be kept in the same cage if it is large enough.
 b. Two magnets either attract or repel each other. (Krifka et al 1995: 35)

In the literature, cases like the ones above are not fully understood yet, and we only have descriptive understanding of them (e.g. see Declerck 1986; Krifka et al 1995: 35). However, we still would like to maintain the analysis of genericity in Krifka et al 1995 and assume that the same analysis also applies to NCs in generic sentences. In (51), we illustrate this analysis with a concrete English example (with irrelevant details omitted).

¹⁵ Some other poorly understood cases include the ones in (3) in which *a*-indefinite ‘a cat’ is accepted in generic sentences, whereas the NC ‘one cat’ is not.

- (3) a. A cat is very self-sufficient. (Heim 1989)
 b. # One cat is very self-sufficient (#generic).

- (51) a. *Three boys* can lift the piano.
 $[[\text{three}]] = \lambda P [\text{three}_{\langle\langle e,t \rangle, \langle e,t \rangle\rangle}(P)]$ $\langle\langle e,t \rangle, \langle e,t \rangle\rangle$
 $[[\text{three boys}]] = \text{three}(\text{boy})$ $\langle e,t \rangle$ ¹⁶
 $[[\text{three boys can lift the piano}]]$
 $= \forall x, s [3(\text{student}_s)(x) \wedge C(x,s)] s' \text{ ACC}(s,s') [\text{lift}_{s'}(x, \text{the piano})]$
 where ACC is the accessibility relation: $\text{ACC}(s, s')$ iff s' is accessible from s ;
 ‘C’ is a contextually salient relation between individuals and situations: $C(x, s)$ iff x is contextually relevant in s (e.g. x is in a situation s in which the piano needs to be lifted) (e.g. see Krifka et al 1995).

So far, we have examined the semantics of numerals and that of NCs; in the rest of this section, we are going to discuss the syntactic status of numerals and the structure of NCs. Two possibilities are available for the syntactic status of numerals: as heads (Ritter 1991; Giusti 1991, 1997; Zamparelli 1995; Simpson 2005, among others) or as phrases (Selkirk 1977; Li 1999; Haegeman & Guéron 1999; Borer 2005; Ionin and Matushansky 2006; Di Sciullo 2012, among others). We adopt the phrasal analysis of numerals for one empirical and one theoretical reason. Empirically, it has been observed that numerals cross-linguistically can undergo coordination to form complex numerals and can be case-marked in some languages; these facts suggest that numerals are phrases rather than heads (e.g. see Zabbal 2005, Ionin and Matushansky 2006, Di Sciullo 2012). Theoretically, the phrasal analysis allows us to provide a recursive and compositional analysis of complex numerals (involving addition or/and multiplication) from simple numerals through conjunction. (e.g. see Ionin and Matushansky 2006 for details).

With the proposed lexical rule of numerals in (45), the structure of NCs (both predicative ones and argumental ones) can be simplified with a uniform syntax across NMLs, as illustrated in (52).

- (52) Structure of NCs in NMLs
 a. $\{n, \{-s, NP\}\}_{\langle e,t \rangle} \rightarrow \text{predicative}$ b. $\{n, \{-s, NP\}\}_{\langle e \rangle} \rightarrow \text{argumental (indefinite)}$
- NumP
 $n_{\langle et, et \rangle}$ $-s_{\langle et, et \rangle}$ $NP_{\langle e, t \rangle}$

NumPf _{\exists}
 $n_{\langle et, e \rangle}$ $-s_{\langle et, et \rangle}$ $NP_{\langle e, t \rangle}$

Again, one may consider a null D analysis of argumental (indefinite) NCs in which a null D associated with the choice function variable merges with the predicative NC in (52a), yielding a DP with an indefinite interpretation, in the same way as a overt definite D

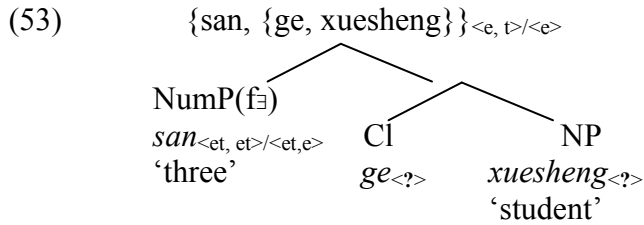
¹⁶ Note that the NC ‘three boys’ here can either remain as a predicate with *Gen* binding it (e.g. see Chierchia 1995), or it can be shifted through the choice function f_{\exists} to an argument with a nonspecific interpretation and being quantified over situations (e.g. see Krifka et al 1995). The semantics of the two cases are compatible with the analysis developed here.

merging with a predicative NC to create a DP with a definite interpretation (e.g. *the two boys*). Nevertheless, as we argued above, a null DP analysis of argumental (indefinite) NCs is inevitably more stipulative than the proposed lexical analysis of numerals.

Summarizing Section 3.2, building on previous analysis of numerals, we showed a general approach of numerals and NCs in NMLs with the goal that this approach should capture the constant behavior of NCs across languages. In section 3.3, we illustrate how this approach applies to CILs.

3.3 NCs in Mandarin: extending the baseline proposal to CILs

As we saw in Section 3.1, the behavior of NCs exhibit exceptional uniformity across languages, so one may wonder whether we can apply the analysis of NCs sketched above to CILs as well. Indeed, it is unavoidable to analyze NCs across languages in a similar way since such an analysis not only explains the cross-linguistically uniform behaviors of NCs, but also it avoids language specific assumptions about NCs. Under such an approach and building on the structure of Mandarin NCs proposed by Li (1997) and Cheng and Sybesma (1999), a rough analysis of NCs in CILs could be the one below.



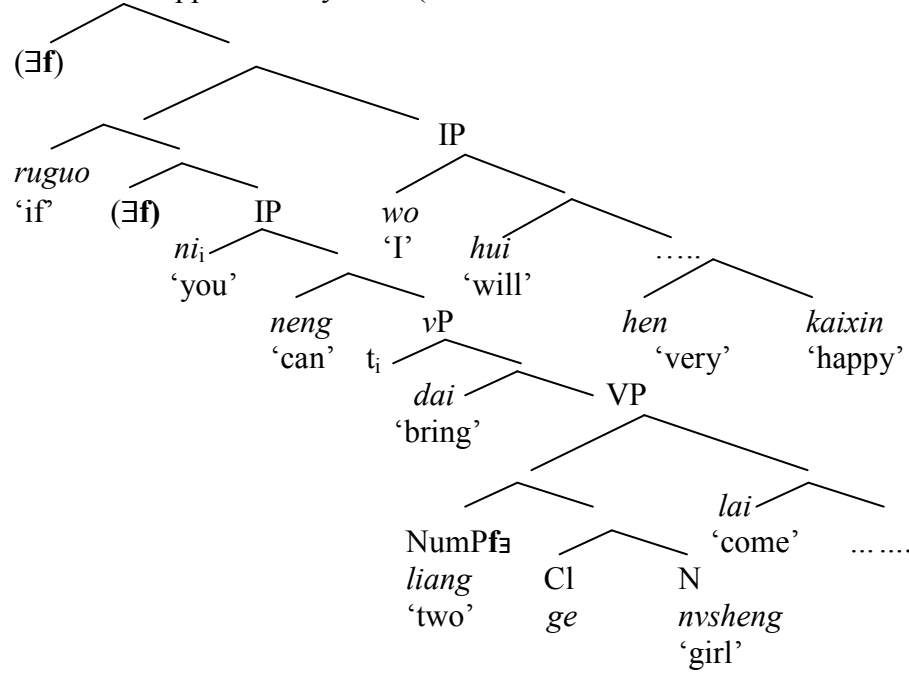
In (53), the behavior of numerals is constant across languages and the output of NCs, i.e. being predicative or argumental, is also the same cross-linguistically. What remains unknown is the semantics of classifiers and that of bare nouns, to be analyzed in details in Section 4. Although we haven't fully understood the internal domain of NCs in CILs yet, this preliminary analysis in (53) is sufficient to explain the predicative use of NCs in CILs as well as their argumental use (i.e. indefinite), as illustrated with two examples from Mandarin below (with the semantics of the unknown parts underlined).

- (54) a. *Zhangsan, Lisi, Wangwu shi san ge xuesheng.*
 Zhangsan Lisi Wangwu be three CI student
 'Zhangsan, Lisi, and Wangwu are three students.'
 b. $[[san \text{ ge } xuesheng]] = [three(\text{ge}(\text{student}))]$ $\langle e, t \rangle$

- (55) a. *ruguo ni neng dai liang ge nvsheng lai wo-de party de-hua,*
 if you can bring two CI girl come my party if
wo hui hen kaixin.
 I will very happy
 'If you can bring two girls to my party, I will be very happy.'

Example (55) (Continued)

- b. $[[liang\ ge\ nvsheng]] = f_{\exists} (two(Cl(girl)))$ <e>
 c. Wide scope reading:
 $\exists f [CH(f) \wedge [bring(you, f(two(Cl(girl)))) \rightarrow happy(I)]]$ [two > if]
 d. Narrow scope reading:
 $[\exists f [CH(f) \wedge bring(you, f(two(Cl(girl))))]] \rightarrow happy(I)$ [if > two]
 e. Existential closure applies at any level (no island constraint violation).



Concerning the generic use of NCs in Mandarin (an example is repeated in (56a)), it has been referred to as ‘the number-denoting interpretation of NCs’ by Li (1998). However, as we have shown in Section 3.1, this type of uses of NCs is better analyzed as generic in contrast with NCs in episodic sentences (e.g. (35)). In particular, we showed the generic interpretation of the nominals is contributed by the *Gen* operator which quantifiers over the whole sentence. The sentence in (56a), hence, can be analyzed in same way as the English counterpart (51), as illustrated in (56b) (with irrelevant details omitted).

- (56) a. san ge nansheng ke-yi tai-qi yi jia gangqin
 three Cl boy can lift-up one Cl piano
 ‘Three boys can lift one piano.’
 b. $\forall x, s [3(Cl(boy_s))(x) \wedge C(x,s)] s' ACC(s,s') \diamond [lift_s(x, one\ piano)]$

Note that, although Mandarin NCs share important properties with those in NMLs at the clausal level, they do exhibit a strong tendency to be interpreted as definite in the sentence initial position in episodic sentences (see Chao 1968; Li & Thompson 1981, among many others). In particular, sentences containing NCs with a nonspecific

interpretation in the sentence initial position are marginal, and a verb *you* ‘exist’ is required to improve the acceptability, as illustrated below.

- (57) ??(you) *san ge xuesheng chile dangao*
exist three CI student ate cake
‘Three students ate the cake.’ (Li 1997: 2, (1b) with slight modification)

To explain this interpretational tendency in Mandarin, some authors assume a covert D for argumental NCs which is subject to syntactic constraints like government (e.g. Li 1998). Nevertheless, some authors did observe a range of well-accepted examples containing indefinite nominal in the sentence initial position in Mandarin (e.g. Huang 1997; Yang 2001; Jiang 2012). A reanalysis of the interpretational tendency in (57) and these well acceptable sentences along the lines consistent with the present D-less approach has been proposed in Yang 2001, E. Tsai 2011 and Jiang 2012. For example, in Yang 2001 and Jiang 2012, it was argued that sentence initial position in Mandarin is topic position, which only allows nominals with a definite or specific interpretation. They then explain the interpretational tendency via a typological account based on Li and Thompson’s 1997 categorization of topic-prominence and subject-prominent languages.¹⁷

Summarizing Section 3, we showed a uniform account of numerals and NCs in which NCs are systematically ambiguous (in the Universal Lexicon) between a predicate modifier and an indefinite variant; this account captures the tendentially universal properties of NCs across languages regardless of whether a language is a CIL or a NML and whether or not a language has Ds. With such an analysis, we look into the internal domain of NCs in Section 4 so as to arrive at the properties that are specific to CILs.

4 The Internal Structure of NCs in CILs

In the internal domain, NCs in CILs share certain similarity with NCs in canonical NMLs.¹⁸ Take English and Mandarin for example, in English, the plural number marking *-s/-es* is obligatory when a numeral (other than ‘one’) combines with a count noun (58a,

¹⁷ Relevant to the present discussion is distribution of ‘plural’ noun phrases in Mandarin, e.g. Li 1999. See Jiang (2012) for a reanalysis of Li’s generalization along lines consistent with the present approach.

¹⁸ ‘Canonical NMLs’ here refers to standard number marking languages such as Romance and Germanic languages where number morphology is obligatory when a numeral combines with any count noun. Canonical NMLs do not include languages like Turkish, Hungarian or Western Armenian, which have number morphology on bare nouns without the occurrence of numerals but *ban* number morphology or optionally allow it on nouns when they appear with a numeral. These languages will be referred to as non-canonical number marking languages. The variation between canonical NMLs and non-canonical NMLs in the literature has been attributed to a different semantics for their numerals (Bale et al 2010), or the different semantics of the functional head determines number morphology on nominals (Sauerland 2003; Sauerland et al 2005; Scontras 2012), or differences with respect to syntactic Agreement (Ritter 1991; Frank 1994). We leave these non-canonical NMLs for future research.

a'). Although Mandarin does not have the same canonical number marking as English, it has classifiers, which are also obligatory when a numeral combines with a noun (58b, b').

- (58) a. *three* *students*_L a'. **three student*
 b. *san* *ge xuesheng* b'. **san xuesheng* (Mandarin)
 three Cl student
 ‘three students’

Setting aside the above similarity, NCs in NMLs and CILs differ significantly in their internal domain. The first difference shows up when the numeral is removed from the nominal domain. When numerals do not occur, phrases in the plural form {-s, N} (in English but not in French) can *freely* occur in argument position (59a), but phrases in the form {Cl, N} in Mandarin cannot, e.g. (20), as repeated in (59b).

- (59) a. i. *Students should study hard.* ii. I taught *farmers, students and workers.*
 b. i. **ge xuesheng yinggai haohao xuexi.* (Mandarin)
 Cl student should good-good study
 Intended: 'A student should study hard.'
 ii. *wo jiao le ge nongmin, ge xuesheng he ge gongren.*
 I teach Asp Cl farmer, Cl student and Cl worker
 Intended: 'I taught a farmer, a student, and a worker.'

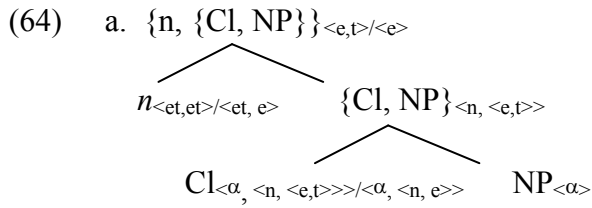
Note that, the difference illustrated in (59) is also one of the two differences between Mandarin and Yi (e.g. (20), (22)), as discussed in Section 2. Another difference between NMLs and CILs emerges when the functional head, i.e. number morphology/classifiers, is absent *together* with the numeral from within the nominal domain. Bare singular nouns in English cannot appear as bare arguments (while in Russian they can) (60a), but in CILs all nouns can *always* merge *directly* with a verb and occur *freely* as bare arguments (60b).

- (60) a. i. *Student should study hard. ii. *I taught student, farmer and worker.
 b. i. xuesheng yinggai haohao xuexi. (Mandarin)
 student should good-good study
 ‘Students should study hard.’
 ii. wo jiao le xuesheng, nongming he gongren.
 I teach Asp student, farmer and worker
 ‘I taught students, farmers and workers.’

Furthermore, number morphology is number specific, i.e. it is either singular or plural (61a), but classifiers are not, i.e. they remain the same regardless of singularity or plurality (61b).

merging with a numeral n . We will adopt the structure in (63) to maintain the connection between the number morphology and the classifier.

Numerals across languages, as we argued, are adjectival functions from properties into quantized properties or functions from properties to entities. When numerals combine with $\{Cl, NP\}$, it follows naturally that $\{Cl, NP\}$ should denote properties, type $\langle e, t \rangle$. What we need to figure out next, then, is the semantics of classifiers and that of bare nouns in $\{Cl, NP\}$. If we assume that nouns in CILs are of type α , the detail of which will be determined below, the type of classifiers might be of $\langle \alpha, \langle n, \langle e, t \rangle \rangle \rangle$ or $\langle \alpha, \langle n, e \rangle \rangle$ (determined by the semantics of the numeral n). Hence, the internal structure and semantics of NCs in CILs could be treated roughly as the one below.



To implement the idea in (64), our first attempt is to assume that nouns in CILs are *property-denoting* of type $\langle e, t \rangle$, the same as those in NMLs. But different from nouns in NMLs, we can assume that those in CILs are undistinguished in their mass versus count denotation (e.g. see Borer 2005: 94; Dayal 2011b). Since their denotations are vague and unspecified, nouns in CILs cannot combine directly with numerals (i.e. numerals need to combine with sets containing atoms/groups). Therefore, it is reasonable to conjecture that some function becomes necessary in such languages to connect numerals with unspecified nouns. Classifiers provide such a function, i.e. they make a noun denotation atomic/countable (e.g. Chierchia 1998b; Borer 2005; Dayal 2011b). This perhaps provides a natural role for classifiers and addresses the issue of why classifiers are obligatory with numerals in CILs. In particular, let us assume that classifiers are relations between numerals and atomized noun denotations along the lines of Krifka 1995. That is to say, numerals can be viewed as one of the two semantic arguments of classifiers. Two examples from Mandarin below illustrate how this implementation works.

(65) First attempt: bare nouns in CILs are property-denoting

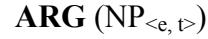
- a. i. *san ge ren* (Mandarin)
 three $Cl_{individual}$ man
 ‘three men’
 ii. $[[ren]] = man$ $\langle e, t \rangle$
 iii. $san_{\langle \langle e, t \rangle, \langle e, t \rangle \rangle} = \lambda P [three_{\langle \langle e, t \rangle, \langle e, t \rangle \rangle} (P)]$, or $\langle \langle e, t \rangle, \langle e, t \rangle \rangle$
 $san_{\langle \langle e, t \rangle, e \rangle} = \lambda P \exists (three_{\langle \langle e, t \rangle, \langle e, t \rangle \rangle} (P))$, $\langle \langle e, t \rangle, e \rangle$
 iv. $[[ge]] = \lambda P \lambda n [n (AT(P))]$ $\langle et, \langle n, e \rangle \rangle$ or $\langle et, \langle n, \langle e, t \rangle \rangle \rangle$
 v. $[[san ge ren]] = three (AT(man)) = [three(AT(man))]$, or $\langle e, t \rangle$
 $= \exists (three(AT(man)))$ $\langle e \rangle$
- b. i. *san bang yingtao* (Mandarin)
 three Cl_{pound} cherries
 ‘three pounds of cherries’
 ii. $[[yingtao]] = cherry$ $\langle e, t \rangle$
 iii. $[[bang]] = \lambda P \lambda n [n (\lambda x [P(x) \wedge pound(x)])]$ $\langle et, \langle n, e \rangle \rangle$ or $\langle et, \langle n, \langle e, t \rangle \rangle \rangle$
 iv. $[[san bang yingtao]] = three (\lambda x [cherry(x) \wedge pound(x)])$
 $= three (\lambda x [cherry(x) \wedge pound(x)])$, or $\langle e, t \rangle$
 $= \exists (three (\lambda x [cherry(x) \wedge pound(x)]))$ $\langle e \rangle$

In (65aiv), *AT* can be understood as a relation between numerals and atomized noun denotations; it atomizes undifferentiated properties and turns them into sets containing singular *atoms* and enables the numeral *san* ‘three’ to combine with them (65av). ‘Atoms’ can be simply defined in the traditional way: as definite singular count nouns (e.g. *the desk*, *the concept*) which are relevant only to the plural domain but not to the mass domain (e.g. Link 1983).¹⁹ The proposed analysis of individual classifiers in (65a) can be extended to other types of classifiers. For instance, in (65b), the measure classifier *bang* ‘pound’ combines with the undifferentiated property *yingtao* ‘cherry’, yielding a property of ‘cherry’ measured in pound (65biii); this newly created properties which contains countable atoms/groups allows the numeral to combine with them (65biv).

Given the non-argumental type of bare nouns in (65), to derive *bare arguments* in Yi and Mandarin (e.g. (13) and (18)), an argumentizing operation becomes necessary for nouns in these languages. Such an operation is not hard to justify: we can assume either a covert D in the syntax (e.g. Borer 2005), or a type-shifting operation in the semantics (e.g. Dayal 2004, 2010), or a combination of the two (e.g. Chierchia 1998). Both covert

¹⁹ Or, assuming that the mass domain and the plural domain both contain atoms and only differ in whether they contain minimal stable atoms (i.e. in the mass domain (e.g. *water*), one cannot define the minimal stable atoms; however, atoms can be well defined in the plural domain (e.g. *apples*)) (Chierchia 2010), the atoms checked by individual classifiers can be defined as the minimal stable ones, and these atoms are still populated the denotations of definite singular count nouns, as in the traditional view. The proposed atomic checking by individual classifiers is compatible with both views of atoms.

(66) a. Syntactic Projection **or** b. Semantic Type-Shifting



- (68) Deriving the interpretations of bare nouns in CILs (*first version*)
- | | | | | | | |
|--------------------------|-----------|-----------|--|-----------|--------------|--------|
| a. <i>ko-lo</i> | <i>gi</i> | <i>o.</i> | b. <i>tshi ko-lo</i> | <i>su</i> | <i>bo o.</i> | (Yi) |
| dinosaur | extinct | SFP | He | dinosaur | look-for | go SFP |
| 'Dinosaurs are extinct.' | | | 'He went to look for dinosaurs/the dinosaur(s).' | | | |
- c. $[[ko-lo]] = \text{dinosaur}$ <e,t>
- d. $[[ARG \cap ko-lo]] = \cap \text{dinosaur}$ <e>
- e. $[[ARG_t ko-lo]] = t \text{ dinosaur}$ <e>
- f. $[[ARG_f \exists ko-lo]] = \text{unavailable, due to (67)}$ <e>
- g. $[[tshi ko-lo su bo o]]$ (via DKP) = $\exists e \exists x [\bigcup \text{dinosaur}(x) \wedge \text{look-for}_w(e, \text{he}, x)]$

Based on the ranking in (67), Yi/Mandarin bare nouns should be first turned into arguments with a kind interpretation via ARG^{\cap} or arguments with a definite interpretation via ARG^{\downarrow} (68d, e). The possibility of turning Yi/Mandarin bare nouns into arguments with an indefinite interpretation via ARG^{\exists} is ruled out because the other two meanings that rank higher (i.e. kind and definite) are already available (68f). The Chierchia-Dayal style analysis correctly predict that bare nouns in both Yi and Mandarin can receive a kind reading and a definite reading but not an indefinite/long-distance scope interpretation (e.g. see (13), (16), (17) and (18)). Note that, in addition to the kind and the definite interpretations, Yi/Mandarin bare nouns also receive a narrow scope existential reading (e.g. (17) and (19iib)). This reading can be derived from the kind interpretation as well, via the operation *Derived Kind Predication* (DKP) as proposed in Chierchia (1992, 1998) (68g). DKP is a sortal adjustment operation; it adjusts predicates when a sentence has a predicate that is primarily object-level with a kind-level argument by introducing a *local* existential quantification over instances of the kind.

To summarize this sub-section, we considered one analysis (65) to carry out the idea of the internal structure of NCs in CILs in (64). We have shown how this analysis explains the three similarities between Yi and Mandarin. In the following sub-section, we will illustrate how this analysis accounts for the differences between these two languages.

4.2 Parameters of CILs: how Yi and Mandarin differ

The differences between Mandarin and Yi, as we propose, can be reduced to two parameters: (i) Yi has a process that derives *bare* CIPs in an unrestricted manner, perhaps via a lexical ‘intransitivization’ process (INTR) for classifiers, but Mandarin doesn’t; (ii) Yi has an lexical D while Mandarin doesn’t. These two parameters may be linked in certain ways, and we remain agnostic as to how they could be related. Regarding these two parameters, the second one [\pm lexical D] is quite straightforward, but the first one [\pm CI_{INTR}] requires some elaboration, which we provide below.

As we have shown in (64), the universal use of classifiers is the ‘transitive’ one, i.e. classifiers have two semantics arguments and obligatorily require a noun and a numeral in their semantics in order to complete a *numeral-taking* classifier phrase {n, {Cl, NP}}. To derive *numeral-less* bare CIPs {Cl, NP} (a variable property attested only in *some* CILs), certain operation becomes necessary to ‘remove’ one of the two arguments (i.e. the numeral) from the classifiers. Below, we consider two such processes.

Mandarin allows bare CIPs in a rather limited distribution (e.g. in a position immediately following a verbal element or a stressed nominal element), as we saw in (20); this can be the result of a process that applies in a *restricted* form, perhaps via a PF *one*-deletion rule along the lines first pursued in Lü 1944 (and elaborated in Chao 1968;

Li 1997; Borer 2005; Huang 2009; Jiang 2012; Li and Feng 2013, among others).²⁰ What this *one*-deletion analysis implies is that bare CIPs in Mandarin are *not* really bare in the syntax, i.e. they have the structure {one, {Cl, NP}}, and eliding *one* during the process of externalization simply results in a *prima facie* bare form of the CIPs {~~one~~, {Cl, NP}}. Semantically, eliding *one* either leads to a long-distance choice function (CF) reading or a GQ reading. As for the semantic difference between the *prima facie* bare CIPs {~~one~~, {Cl, NP}} and NC {one, {Cl, NP}}, it lies in that the former cannot be used in contrastive or focused environments where the information conveyed by the numeral *one* is stressed/important (e.g. Lü 1944, Li and Bisang 2011).²¹ This difference can be viewed as the semantic condition for eliding *one*. Except for this difference, Mandarin numeral-less CIPs and numeral-taking CIPs do share two main interpretations: the narrow scope existential reading (i.e. the GQ reading) (e.g. Cheng and Sybesma 1999, Li and Bisang 2011, Jiang 2012) and the long-distance scope reading (i.e. the choice function reading) (e.g. Jiang 2012). We illustrate this *one*-deletion analysis of Mandarin bare CIPs in (69).

- (69) *One*-deletion, deriving bare CIPs in Mandarin (applied in a restricted way)
- a. *mei ge youke dou bei (yi) ge nvhai-r huyou mai le yi bu shouji.*
 every CI visitor all Pass one CI girl hoodwink buy Asp one CI cell-phone
 ‘Every tourist was hoodwinked to buy a phone by a girl.’ [a girl > ∀], [∀ > a girl]
 (Jiang 2012:184, ex (153))
- b. Syntax
- {yi, {ge, nvhai-r}}
 NumP
yi
-
- c. PF
 [~~yi~~ ge nvhai-r]
- $\begin{array}{c} \text{CI} \quad \text{NP} \\ \text{ge} \quad \text{nvhai-r} \end{array}$
- d. Semantic condition for eliding *yi* ‘one’: when the numeral *one* information is not important/stressed/focused.
- e. Interpretations of bare CIPs: [~~one~~ Cl N] = [one Cl N]
- i. [[~~one~~ Cl N]] = $\lambda P \exists f [P (f (\lambda x \text{ one } (P)(x)))]$ elided *one* → GQ reading
- ii. [[~~one~~ Cl N]] = $f_{\exists} (\lambda x [\text{one } (AT (P))(x)])$ elided *one* → CF reading

²⁰ Both Cheng and Sybesma 1999 and Li and Bisang 2011 provide arguments against the *one*-deletion view of Mandarin bare CIPs first proposed in Lü 1944, but see Li 1997, Borer 2005, Huang 2009, Jiang 2012, and Li and Feng 2013 for arguments in favor of the *one*-deletion approach.

²¹ Two examples to illustrate the difference between numeral-less CIPs and numeral-taking CIPs in Mandarin are given in (4) in which the numeral *one* cannot be omitted.

(4) a. *you *(yi) ge xuesheng, wu ge chefu.*
 have one CI student five CI carter
 ‘There is a student and five carters’
 (Lü 1944: 167, with slight modification)

b. *wo hua le *(yi) ge xiaoshi chifan*
 I spend Perf one CI hour eat-meal
 ‘I spent one hour eating a meal.’
 (Li and Bisang 2012: 345, ex (18))

This difference has a straightforward explanation, namely that in contrastive/focused environment, contrastive/focal stress is required; however, null elements are destressed.

The *one*-deletion analysis also logically explains why Mandarin bare CIPs {Cl, NP} can only be interpreted as singular but rather ‘two’, ‘three’, or ‘some’: as it is ‘one’ that is deleted from [*one* Cl-N], the [Cl- N] phrase can only be interpreted as singular one.²² With regard to the distribution of bare CIPs in Mandarin (e.g. (20)), it could be viewed as the result of either some phonological restrictions (e.g. Lü 1944, Jiang 2012; Li and Feng 2013) or some syntactic restrictions like government (e.g. Cheng and Sybesma 1999, Huang 2009, Li and Bisang 2012).

Turning to Yi, it allows *unrestricted* bare CIPs (i.e. bare CIPs can appear in any argument position) (e.g. (22)); this could be due to a process that derives them in an unrestricted manner, perhaps via a *lexical* ‘intransitivization’ process (INTR) for classifiers, as illustrated in (70).

- (70) Lexical Rule *INTR*, deriving bare CIPs in Yi (applied in unrestricted way)²³
- a. $Cl = \lambda P \lambda n [n(AT(P))]$ or $= \lambda P \lambda n [f_{\exists}(n(AT(P)))]$
 - b. if $\alpha \in Cl$, $\lambda P \lambda x [Cl(One)(P)(x)] \in Cl_{INTR}$
 - c. $Cl_{INTR} = \lambda P \lambda x Cl_{INTR}(One)(P)(x) = \lambda P \lambda x [One(AT(P))(x)]$
 - d.
$$\begin{array}{ccc} & \text{CIP}_{\langle e, t \rangle / \langle e \rangle} & \\ & \swarrow \quad \searrow & \\ \text{NumP} & & \text{CIP} \\ n_{\langle et, et \rangle / \langle et, e \rangle} & & \end{array} \xrightarrow{\text{INTR}} \begin{array}{ccc} & \text{CIP}_{\langle et, \langle n, et \rangle \rangle / \langle et, \langle n, e \rangle \rangle} & \\ & \swarrow \quad \searrow & \\ \text{NP}_{\langle e, t \rangle} & & \text{Cl}_{\langle et, \langle n, et \rangle \rangle / \langle et, \langle n, e \rangle \rangle} \end{array}$$
 - e. $Cl_{INTR}P = \lambda x [One(AT(P))] \langle e, t \rangle$

The lexical rule in (70) removes one of the semantic arguments, i.e. the numeral *n*, from the classifiers, yielding ‘intransitive’ classifiers. The availability of INTR is to be considered as a parameter [$\pm Cl_{INTR}$], which a language may or may not have. The lexical rule accounts for the variable property of CILs, explaining why some languages permit bare CLPs (e.g. Yi) but some don’t (e.g. Japanese, Korean).

After an intransitivized classifier merges with a noun, the structure in (70iie) creates property-denoting nominals. In order to achieve the freely argumental use of the classifier-noun phrase {Cl, NP} in Yi (e.g. (22)), this property-denoting nominal can be turned into an argument via whatever device available in the language. For the purpose of simplification, we still use the Chierchia-Dayal style account to illustrate how property-denoting {Cl, NP} phrases are argumentized, as in (71).

²² We can take the view that the numeral *one* in bare CIPs is either realized in a phonetically null form or is deleted at PF; the difference between the two does not make a difference for the purpose of our discussion.

²³ It might be appealing to propose another analysis for bare CIPs in Yi: unlike Mandarin, Yi allows *one*-deletion everywhere given that ‘one’ is sandwiched between N and CL in Yi, that is, what comes immediately before the nominal in Yi does not matter, because that element is not adjacent to the numeral ‘one’. However, such an analysis, although also applicable to Bangla, will not apply to other CILs like Cantonese, i.e. its numerals precede both N and Cl, but bare CIPs are also freely allowed.

- (71) Argumentizing $\{Cl, NP\}$ in Yi
- a. $[[NP\ Cl\ su]] = \iota \lambda x [one(AT(P))]$
 - b. If $\{Cl, NP\}$ is argumentized covertly via ARG, it can only be indefinite since:
 - i. $ARG_{\iota} \{Cl, NP\} = \text{blocked by the presence of } su$
 - ii. $ARG^{\cap} \{Cl, NP\} = \text{undefined for singular properties}$
 - iii. $ARG_{\exists} \{Cl, NP\} = \text{indefinite}$

In (71a), the overt definite determiner *su* can turn $\{Cl, NP\}$ into arguments with a definite interpretation. In addition to overt operations like lexical determiners, $\{Cl, NP\}$ can also be turned into arguments covertly via ARG (e.g. (66)). In principle, ARG can turn $\{Cl, NP\}$ into arguments with a definite or a kind interpretation due to *Ranking of Meanings* in (67). Nevertheless, the two covert options, i.e. ARG_{ι} and ARG^{\cap} , are unavailable for the following reasons. First, the presence of an overt definite article *su* in Yi blocks the possibility of turning Yi bare ClPs covertly into arguments with a definite reading (the *Blocking Principle*, see Chierchia 1998) (71bi). Second, turning $\{Cl, NP\}$ to kinds is undefined because the semantics of singularity of $\{Cl, NP\}$ (e.g. (65aiv)) clashes with the conceptual notion of a kind which corresponds to the plurality of all instances of the property (see Dayal 1992, Chierchia 1998) (71bii). Accordingly, the only possibility left, according to *Ranking of Meanings*, is to turn $\{Cl, NP\}$ covertly into arguments with an indefinite reading via ARG_{\exists} (71biii). This indeed is what happens in Yi: its bare ClPs are indefinite only (e.g. (22)).

As we can see, the first analysis of the internal semantics and structure of NCs in CILs outlined in (65), which assumes that their nouns are mass-count undistinguished *properties*, has very promising consequences. For instance, it not only justifies the status of classifiers and accounts for the distribution and interpretation of NCs and bare arguments in Mandarin and Yi in a uniform way, but it also reduces differences between them to two main parameters: $[\pm Cl_{INTR}]$ in the lexicon and $[\pm D]$ in the syntax. As for the third parameter $[\pm PF\ one\text{-}deletion]$ in the process of externalization, it is only relevant to CILs like Mandarin which do not have intransitivized classifiers. While this is promising, this analysis faces two main problems which we will address in the next sub-section.

4.3 Problems with the first analysis and a new proposal

The first problem of the analysis in (65) is that if bare nouns in CILs have the *same* type as those in NMLs (i.e. *properties*), one would expect there to be a CIL in which nouns *cannot* occur bare as arguments and overt Ds are obligatory, just like there are NMLs with such properties, e.g. French. Nevertheless, analogues of French are *not* attested among CILs.²⁴ The second problem faced by the first analysis in (65) is that if Ds are

²⁴ It is true that so far we only observe one language, namely French, which obligatorily projects overt Ds above *all* bare nouns (bare singulars, bare plurals, and mass nouns). If languages disallowing bare

universally property taking (of, say, type $\langle e, t, e \rangle$), nothing prevents Ds from combining with bare nouns in CILs (as they in fact do in NMLs). However, in Yi this does not happen.²⁵ The definite determiner *su* simply cannot combine with bare nouns (28), contrary to the prediction, although it can combine freely with higher nominal structures (i.e. NCs and bare CIPs (23)).

These two problems lead us to an alternative analysis in which nouns in CILs are *kind-referring* (i.e. entities, see Carlson 1977b for reasons why kinds are entities). The purpose of classifiers, then, can be viewed as to relate kind-denoting nouns with numerals, along the lines developed in Krifka 1995 and Chierchia 1998. Everything else proposed in Section 4.1 and 4.2 remains the same. This second analysis is given in (72), as illustrated with the same examples from Mandarin.

(72) Second analysis: bare nouns in CILs are kind-denoting

- a. i. *san ge ren*
 three $Cl_{\text{individual}}$ man
 ‘three men’
 ii. $[[ren]] = \text{man}$ $\langle e \rangle$
 iii. $\text{three}_{\langle e, t, \langle e, t \rangle \rangle} = \lambda P [\text{three}_{\langle e, t, \langle e, t \rangle \rangle} (P)]$, or $\langle \langle e, t \rangle, \langle e, t \rangle \rangle$
 $\text{three}_{\langle e, t, e \rangle} = \lambda P [f_{\exists} (\text{three}_{\langle e, t, \langle e, t \rangle \rangle} (P))]$, $\langle \langle e, t \rangle, e \rangle$
 iv. $[[ge]] = \lambda k \lambda n [n (AT(^k))]$ $\langle e, \langle n, e \rangle \rangle$ or $\langle e, \langle n, \langle e, t \rangle \rangle \rangle$
 v. $[[\text{three ge ren}]] = \text{three} (AT(^{\cup} \text{man})) = [\text{three} (AT(^{\cup} \text{man}))]$, or $\langle e, t \rangle$
 $= f_{\exists} (\text{three} (AT(^{\cup} \text{man})))$ $\langle e \rangle$
- b. i. *san bang yingtao*
 three Cl_{pound} cherries
 ‘three pounds of cherries’
 ii. $[[yingtao]] = \text{cherry}$ $\langle e \rangle$
 iii. $[[bang]] = \lambda k \lambda n [n (\lambda x [^{\cup} k(x) \wedge \text{pound}(x)])]$ $\langle e, \langle n, e \rangle \rangle$ or $\langle e, \langle n, \langle e, t \rangle \rangle \rangle$
 iv. $[[\text{san bang yingtao}]] = \text{three} (\lambda^{\cup} \text{cherry}(x) \wedge \text{pound}(x))$
 $= [\text{three} (^{\cup} \text{cherry}(x) \wedge \text{pound}(x))]$, or $\langle e, t \rangle$
 $= f_{\exists} (\text{three} (^{\cup} \text{cherry}(x) \wedge \text{pound}(x)))$ $\langle e \rangle$

arguments to the degree that French does are already rare, very little can be concluded from the fact that a language which is a combination of Mandarin and French does not exist. Nevertheless, it is crucial that such a gap should be principled rather than accidental (Amy Rose Deal p.c.). In other words, there should, in principle, be other languages like French which disallow bare arguments (i.e. the requirement imposed by French is actually not so rare), but we should *not* expect these languages to be CILs as well. The first attempt in (65) wrongly predicts the existence of such a CIL. In addition, there isn’t a CIL like English that requires articles in certain context (e.g. *the dog*); the treatment of bare nouns in (65) also wrongly predicts the existence of such a CIL.

²⁵ Bangla, another CIL that has been argued to have Ds (e.g. Dayal 2010) also does not allow its bare nouns to be turned into arguments with a definite interpretation via Ds (see Dayal 2010 et seq; Jiang 2012).

In (72), nouns are inherently argumental (e.g. (72aii)). This analysis has three immediate consequences. First, in CILs, bare nouns are kinds ($\langle e \rangle$), so Ds (if they are universally functions from properties to entities) will be unable to combine with bare nouns for type theoretical reasons. It, hence, explains why the definite determiner *su* cannot combine with bare nouns in Yi, as repeated in (73a) and demonstrated in (73b).

- (73) a. **si-hni su dzi ndza*. (Yi)
 girl SU very beautiful
 b. \rightarrow CRASH! composition cannot proceed, type-mismatch
 (argumental) $NP_{\langle e \rangle}$ $D_{\langle \langle e, t \rangle, e \rangle}$ (look for properties)

Regarding Mandarin, since its nouns and NCs are both inherently argumental, it is no longer necessary, nor even possible to assume a covert functional category D in order to account for the behavior of its nominal arguments. This is arguably a simpler and more elegant analysis of Mandarin nominals since it avoids stipulating the presence of invisible projections that otherwise have no overt manifestation in this language. Hence, nominal arguments in Mandarin, according to our analysis, are all D-less (74).

- (74) Mandarin bare nouns, NCs, and bare CIPs: all D-less
- i. Bare nouns
 $NP_{\langle e \rangle} \rightarrow$ argumental
 |
 N
- ii. NCs
 a. $\{n, \{Cl, NP\}\}_{\langle e, t \rangle} \rightarrow$ predicative
 NumP $Cl_{\langle e, \langle n, \langle e, t \rangle \rangle}$ $NP_{\langle e \rangle}$
 $n_{\langle et, et \rangle}$
 b. $\{n, \{Cl, NP\}\}_{\langle e \rangle} \rightarrow$ argumental (indefinite)
 NumP \exists $Cl_{\langle e, \langle n, e \rangle \rangle}$ $NP_{\langle e \rangle}$
 $n_{\langle et, e \rangle}$
- iii. Bare CIPs (not bare in the syntax)
 a. $\{one, \{Cl, NP\}\} \rightarrow$
 NumP Cl NP
 one
- b. [~~one~~ Cl NP] (PF)

It is worth pointing out that Chierchia (1998) speculates that since CILs do not need determiners to argumentize their nouns, they will never develop lexical determiners like *the* in English because of economy considerations. Nevertheless, this speculation is incorrect and this line of reasoning is also faulty. As we saw, Yi does have a lexical determiner, *su*. This fact is indeed expected, contra Chierchia (1998)'s speculation: if some higher nominal projections above the bare noun, namely bare CIPs and NCs, are of

a predicative type $\langle e, t \rangle$, lexical determiners become relevant and may well develop to turn them into arguments.²⁶ Hence, the second immediate consequence of the proposed analysis in (72) is that if a CIL develops a determiner, it will be able to apply only at the level of NCs, which are property-denoting, but *not* at the level of bare nouns, which is kind-denoting, given type theoretic considerations. If an intermediate projection between NCs and bare nouns that is of type $\langle e, t \rangle$, namely bare CLPs, is available, determiners can apply at this level as well. This is precisely what happens in Yi, as repeated in (75) and illustrated in (76).

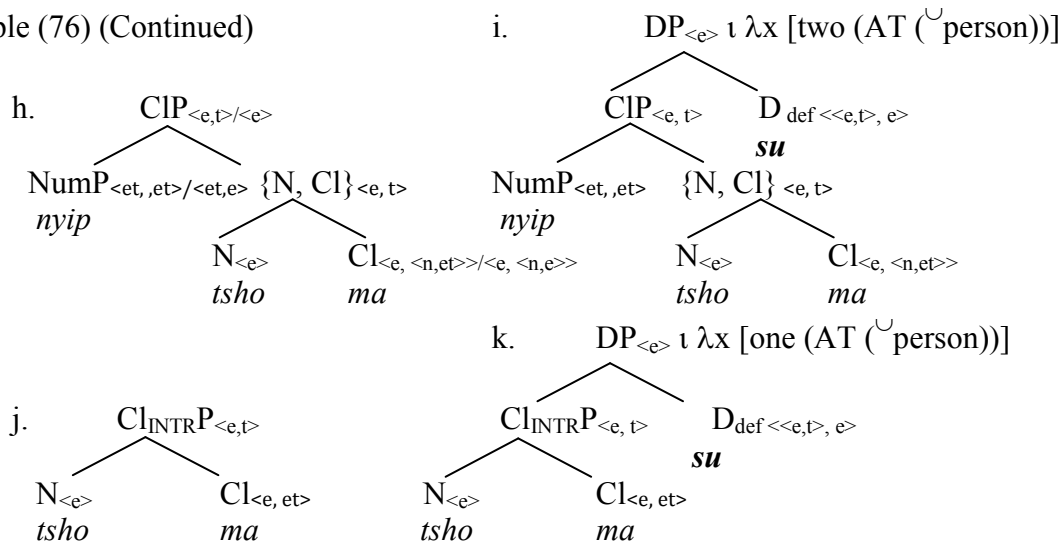
- (75) a. *tsho nyip ma* b. *tsho nyip ma su* c. *tsho ma* d. *tsho ma su*
 person two CL person two CL SU person CL person CL Su
 ‘two persons’ ‘the two persons’ ‘a person’ ‘the person’

- (76) Yi NCs, bare nouns, and bare CLPs: D only applies to bare CLPs and above²⁷
- a. Numeral $\langle\langle e, t \rangle, \langle e, t \rangle\rangle = \lambda P [n (P)]$ or $\langle\langle e, t \rangle, \langle e, t \rangle\rangle$
 $= \lambda P [f_{\exists} (n (P))]$ $\langle\langle e, t \rangle, e \rangle$
 - b. $[[tsho]] = \text{person}$ $\langle e \rangle$
 - c. $[[ma]] = \lambda k \lambda n [n (AT(^{\cup}k))]$ or $\langle e, \langle n, \langle e, t \rangle \rangle \rangle$
 $= \lambda k [\text{one} (AT(^{\cup}k))]$ $\langle e, \langle e, t \rangle \rangle$
 - d. $[[tsho nyip ma]] = \text{two} (AT (^{\cup} \text{person}))$ or $\langle e, t \rangle$
 $= f_{\exists} (\text{two} (AT (^{\cup} \text{person})))$ $\langle e \rangle$
 - e. $[[tsho ma]] = \text{one} (AT (^{\cup} \text{person}))$ $\langle e, t \rangle$
 - f. $[[tsho nyip ma su]] = \iota \lambda x [\text{two} (AT (^{\cup} \text{person}))]$ $\langle e \rangle$
 - g. $[[tsho ma su]] = \iota \lambda x [\text{one} (AT (^{\cup} \text{person}))]$ $\langle e \rangle$

²⁶ Although Chierchia (1998) appeals to an economy condition to rule out the presence of Ds in CILs, such an appeal is too weak since no mention is made regarding the nature of the economy condition that is supposed to be employed. At most it can be seen as a loose appeal to economy conditions because if it were true, CILs would have only one way to express definiteness, through bare nouns, and any other way of marking definiteness, say via type-shifting operations, would have to be regarded as uneconomical.

²⁷ The structure of NCs in (76i) does not provide the linear order [N-Num-CL] in Yi. In the proposed structure, a classifier is in the head-final position. The motivation is that Yi is a head-final language in both the clausal and nominal domains. The headedness in Yi suggests that classifiers should also appear in the head position following nouns. In addition, we assume that numerals left-adjoin to CLPs in Yi. As a consequence, numerals and classifiers are not adjacent to form the Num+CL sequence because nouns separate them. In order to obtain the Num+CL sequence, two options are available. One is that after the structure in (76i) is sent to PF, some rearrangement takes place there which fixes the linear order. The other option is that the noun undergoes movement to a higher position to give rise to the Num+CL sequence.

Example (76) (Continued)



The third immediate consequence of the analysis in (72) is that it correctly predicts that an obligatory CIL in which nouns cannot occur bare as arguments (the analogue of French in the domain of CILs) should not exist. Given that bare nouns in CILs are always argumental, we should expect that any CIL always allows bare nouns to merge directly with verbs and occur freely in argument position.

In addition to the three immediate consequences above, we expect a series of further consequences to follow from the approach in (72), which we will address in Section 5.

5 Further consequences and future vistas

5.1 Definites and demonstratives

The first additional consequence concerns the interpretation of bare nouns in CILs, especially the definite one. We illustrate this with examples from Mandarin in (77a-c).

(77) Deriving the interpretations of bare nouns in CILs (*final version*)

- a. *konglong miejue le*. b. *Mali pa gou*. c. *wo kanjian gou le*. (Mandarin)
dinosaur extinct Asp Mali afraid dog I see dog Asp
‘Dinosaurs are extinct.’ ‘Mali is afraid of dogs.’ ‘I saw dogs/the dog(s).’
- d. $[[\textit{konglong miejue le}]] = \text{extinct}(\text{dinosaur})$
- e. $[[\textit{Mali pa gou}]] = \text{Gen } x [\cup \text{dog}(x)] [\text{be afraid of}(\text{Mali}, x)]$
- f. $[[\textit{wo kan jian gou le}]] (\text{via DKP}) = \exists e \exists x [\cup \text{dog}(x) \wedge \text{see}_w(e, I, x)]$
- g. $[[\textit{gou}]] = [\text{dog}_{\langle e \rangle}]s = \lambda s' \text{dogs}'(s) = \text{the maximal member instantiating dog-kind in a situation } s$

Given that bare nouns are argumental (in this case, kind-referring), their kind-level readings and generic readings then come for free (77d, e); their narrow scope existential reading can be derived from their kind readings via DKP (e.g. Chierchia 1998) (77f). The definite reading for bare nouns in Mandarin and Yi, as we propose, might be obtained by plugging into the kind a situation variable, provided by the context (77g) (see also Trinh 2010 and Dayal 2011a, to appear for proposals along the same lines). From here on, let *Situation Restriction* (SR) refer to this strategy, as shown in (78).

(78) *Situation Restriction* (SR)

$[N_{\langle e \rangle}]s \rightarrow [N_{\langle e_i \rangle}] = \text{the maximal member instantiating } N\text{-kind in a situation } s$

Since kinds are functions from situations to (possibly plural) individuals (e.g. see Chierchia 1998b: 349), by applying them to a ‘resource’ situation (an analogue of Domain Selection, e.g. see von Stechow 1994; Elbourne 2001) we get a situation restricted individual e_i . SR, then, can be viewed as a function from kinds to situation restricted kinds, type $\langle e, e_i \rangle$. It restricts a kind to a specific situation and obtains the *maximal* members in that situation provided by the context. In a sense, SR is analogous to the iota operator (ι), the canonical Frege-Russell definition of which is ‘the largest member of X if there is one (else, undefined)’ (c.f. Chierchia 1998b: 346 and see Sharvy 1980 for the same analysis for English definite determiner). (One might also note that reference to the maximal sum in a situation here plays a role similar to reference to stages in Carlson’s 1977b theory.)²⁸ In a sense, what (77) and (78) show is that the different interpretations of bare nouns in CILs can be derived from their kind reading, including the definite one.

The next consequence is about demonstratives (Dems). As we have shown in Section 1, Dems in Yi, like the determiner *su*, combine only with higher nominal projections (CLPs and above) and never combine with bare nouns (4), as repeated in (79a-c). In contrast, Dems in Mandarin can freely combine with bare nouns as well as higher projections (5), as repeated in (79a’-c’).

- | | | | | |
|------|---|--|--|------------|
| (79) | a. <i>tsho</i> <u><i>tshɿ</i></u> <i>ma</i>
man this Cl
‘the man’ | b. <i>tsho</i> <u><i>tshɿ</i></u> <i>nyip ma</i>
man this two Cl
‘these two men’ | c. * <i>tsho</i> <u><i>tshɿ</i></u>
man this | (Yi) |
| | a’. <u><i>zhe</i></u> <i>ge ren</i>
this Cl man
‘this man’ | b’. <u><i>zhe</i></u> <i>liang ge ren</i>
this two Cl man
‘these two men’ | c’. <u><i>zhe</i></u> <i>ren</i>
this man
‘this man’ | (Mandarin) |

²⁸Note that there could be a number of ways of obtaining definiteness from kinds. SR is one such way, so is a new type of iota of type $\langle ek, e \rangle$ which directly gives rise to the definiteness of kind-referring bare nouns. However, they are both the same idea, and the ‘new iota’ will be undistinguishable from the situation restriction that we proposed.

The different behaviors of Dems in Yi and Mandarin, as we propose, reveal a difference between definite articles and Dems: the former universally combines with properties and return entities, type $\langle\langle e, t \rangle, e\rangle$; whereas the latter can be obtained either from properties (e.g. (79a, b) and (79a', b')) or from kinds (e.g. (79c')) via the help of indexically individuated situations, as illustrated in (80a) and (80b) respectively.

- (80) $\text{this}_n/\text{that}_n$ denotes a (distal) situation that the speaker is pointing at
 a. $\text{this}_n/\text{that}_n (P)(x) = P(x) \wedge x$ is in $\text{this}_n/\text{that}_n$
 b. $K_{\text{this}_n/\text{that}_n} = K(\text{this}_n/\text{that}_n) =$ the manifestation of K in $\text{this}_n/\text{that}_n$

In particular, the same strategy to derive definites from kinds via SR, plus an indexical component, can be exploited to obtain the definite reading for $\{\text{Dem}, \text{NP}\}$ in Mandarin. For example, if the cat-kind c is a function from situations s to the maximal entity that instantiates c in s , *zhe mao* ‘this cat’ in (81a) can be represented as $c(\text{this}_n)$, where this_n denotes the (distal) situation the speaker is pointing at (81b).

- (81) a. *wo xihuan #(zhe) mao bu xihuan #(zhe) mao.*
 I like this cat not like this cat
 ‘I like this cat but don’t like this cat.’
 b. $[[\text{zhe mao}]] = [\text{cat}_{\langle e, k \rangle}](\text{this}_n) =$ the manifestation of c in this_n $\langle e \rangle$

What (81) shows is that Dems, *unlike* definite articles, can apply at different levels in nominal structures, as illustrated in (82) (with optional phrases/heads in parenthesis).

- (82) a. Mandarin $\langle e \rangle$ b. Yi²⁹ $\langle e \rangle$
-

The analyses of definite bare nouns and demonstratives above not only reveal a fundamental difference between definite determiners and demonstratives, but also show that whether or not a CIL has overt Ds does not change the availability of the definite reading of its bare nouns.

²⁹ The linear order can be fixed either in the narrow syntax, e.g. via moving NP to the nominal initial position, or in the PF, e.g. some phonological rules require nouns to be pronounced in the nominal initial position in Yi.

In addition to the above two further consequences, our analysis of CILs has one additional consequence concerning language variation in CILs with respect to the definite reading of their bare nouns, which will be discussed in Section 5.2.

5.2. Language variation regarding definite readings of bare nouns

Although CILs with ‘real’ determiners are rare, there are several CILs where definiteness is somehow overtly marked. Particularly prominent cases are Cantonese (to pick another Sino-Tibetan language) and Bangla (to pick a typologically unrelated Indo-European language). As is well known, in Cantonese, bare CIPs constitute the canonical way to convey the definite interpretation of a nominal (e.g. Cheng and Sybesma 1999) (83a).

- (83) a. *bzek gau zungji sek juk.* (Cantonese)
 CL dog like eat meat
 ‘The dog likes to eat meat.’
 (not: dogs like eating meat) (Cheng and Sybesma 1999: 511 (6a))
- b. *boi ta* (Bangla)
 book CL
 ‘the book’ (Dayal 2011b: 3, (12a’))

Similarly, as illustrated in (83b), definiteness is also marked by bare CIPs (with inverted word order with respect to the basic one) in Bangla (e.g. Bhattacharya 1999a, b; Dayal 2010, 2011b, to appear).³⁰ Interestingly, in both languages, bare nouns generally do not receive a definite interpretation (see Cheng and Sybesma 1999 for Cantonese and Dayal 2011b, to appear for Bangla), as illustrated in (84).

- (84) a. *Wufei jam-jyun *(wun) tong la.* (Cantonese)
 Wufei drink-finish CL soup SFP
 ‘Wufei finished drinking the soup.’ (Cheng and Sybesma 1999: 510 (3b))
- b. *Ami room-ta ek-ta biral key dekhe chhi,* (Bangla)
 I room-CL one- CL cat key see Past
 biral #(ta) amake kamor dilo
 cat CL me-Acc bite have
 ‘I saw one cat in the room. The cat has bitten me.’

What this suggests is the following generalization: definiteness marking of the type found in Cantonese/Bangla is seemingly incompatible with the definite use of the bare noun, whereas the presence of the overt definite article in Yi seems to be compatible with a

³⁰ The order [NP-Num-CL] might not be the basic order of Bangla numeral-containing phrases (see Bhattacharya 1999a, b and Dayal 2010, 2011b, to appear for analyses of Bangla nominal structures).

definite interpretation. While we may not have a final account of this phenomenon, we sketch out two tentative explanations below.

The first account assumes that *SR* that shifts kind-referring bare nouns to definites by plugging in a contextually salient situation $[N]_{SR}$ (as proposed in Section 5.1) is *not* universally available in CILs. If a CIL does not have *SR*, they must have some other way of forming definites (e.g. for expressibility). Arguably, this is the situation that obtains in Cantonese/Bangla; the way to make things definite in these two languages is to use a classifier and a noun $\{Cl, NP\}$ (83). On the other hand, if a CIL *does* have *SR*, it may (e.g. Yi) or may not (e.g. Mandarin) have a definite *D* which applies to property-denoting nominals of type $\langle e, t \rangle$. The reason is because D_{def} and *SR* are different strategies that apply at different levels in the nominal domain, so their coexistence is not conflicting, and they should not compete with each other (85).

(85) $[\alpha_{\langle ek \rangle}]_{SR}$ *versus* $D_{def} [Cl \alpha]_{\langle e, t \rangle}$

No competition:

- i) The type of the *SR* and the type of D_{def} differ: $\langle e_k, e \rangle$ vs $\langle \langle e, t \rangle, e \rangle$
- ii) The levels that they apply to differ: $[\alpha]$ vs $[Cl \alpha]$

The account assuming that *SR* is not universally available predicts that the definite interpretation of bare nouns will be banned in CILs such as Cantonese and Bangla. This account does explain cases like the ones in (84) well; however, the prediction is too ‘strict’ in the sense that it will exclude some accepted cases. As observed in Wu and Bodomo 2009 and Simpson et al 2011, if salient contexts are available, bare nouns in Cantonese and Bangla can be used refer to the unique entity in the context, i.e. they can be definite. According to the above account, cases that allow bare nouns to express definiteness in these two languages, would call for some other explanation.

Alternatively, *SR* can be maintained to be universally available, and some independent reasons could be assumed to block *SR* from applying to bare nouns in Cantonese/Bangla (e.g. some revised version of *Blocking Principle* in Chierchia 1998). Along this line, we propose a ‘soft competition’ between the two covert operations: *SR* in the semantics and ARG in the syntax/semantics. This ‘soft competition’ states that the covert operation *SR* for α will be blocked if ARG can covertly shift $[\beta \alpha]$, a *larger* property-denoting phrase (i.e. $[NP Cl]$) to definites, as formulated in (86).

(86) A ‘soft competition’ exists between *two covert* operations

For any covert type shifting operation in semantics τ and any α

* $\tau(\alpha)$

If there exists a β which turns α to properties such that for any set α in its domain

$[ARG [\beta \alpha]] = \tau(\alpha)$

A crucial element in this soft competition is that the two covert argumentizing operations are not of the same type: *SR* applies to kinds, type $\langle e_k, ei \rangle$; whereas *ARG* applies to properties of type $\langle \langle e, t \rangle, e \rangle$. ‘Soft’ here means that *ARG* does not block *SR* completely. The reason to consider the competition ‘soft’ is based on the difference between the two *covert* operations and the phrases that these two operations apply to (i.e. bare nouns and bare CIPs). To be concrete, to place operations of different types which apply to phrases at different structural levels in a competition is not fair to begin with, thus providing reasons for the ‘soft’ nature of such a competition. As a ‘soft constraint’, we could expect that the bare CIP in Cantonese/Bangla is the preferred/default structure to express definiteness and that their bare nouns still can express definiteness in some circumstances. This, indeed, is the case.

As for Yi, the soft competition in (86) is not applicable because this language has an overt definite determiner *su* that overtly shifts properties into arguments and does not use *ARG* for definiteness. No competition should be placed between an overt operation and a covert operation which apply at different levels (bare nouns and bare CIPs) and are of different types ($\langle e_k, e \rangle$ and $\langle \langle e, t \rangle, e \rangle$), as illustrated in (87).

(87) No competition between *overt* operation and *covert* operation at *different* levels

$[\alpha_{\langle e_k \rangle}]_{SR}$ *versus* $[_{DP} \text{determiner}_{def} [\beta \alpha]_{\langle e, t \rangle}]$

No competition:

- a. overt operations via a determiner do not compete with covert ones at different levels
- b. The levels that the two operations apply to differ: $[\alpha]$ vs $[\beta \alpha]$
- c. The type of the *SR* and the type of definite determiners differ: $\langle e_k, e \rangle$ vs $\langle \langle e, t \rangle, e \rangle$

We have seen three further consequences that follow from the proposed analysis of nominal arguments in CILs. In the following sub-section, we will discuss possible typological investigations that we leave for the future research; we speculate that different settings of the two main parameters that differentiate Mandarin and Yi (i.e. $[\pm D, \pm Cl_{INTR}]$) (as proposed in Section 4.2) predict further language types, which are indeed attested.

5.3 Typological investigation for the future

As we have shown in Section 4.2, the two parameters $[\pm D, \pm Cl_{INTR}]$ captured two types of CILs. The first type lacks the functional category D in the syntax as well as the lexical intransitivization rule for classifiers (i.e. $[-D, -Cl_{INTR}]$ CILs); this type is represented by Mandarin in which unrestricted bare CIPs are not admitted (88i). The second type of CILs that was discussed has both the functional category D and the lexical

intransitivization rule for classifiers (i.e. [+D, +Cl_{INTR}] CILs), as represented by Yi (88ii). In this type of CILs, bare CIPs are freely allowed in argument position, and Ds are only detected at the NC level and the bare CIP level (if allowed). Since nothing in the theory predicts a one-to-one correspondence between the presence of D and the lexical rule of classifiers, our theory predicts that two other types of CILs should be attested: one that does not have Ds but has the lexical rule for classifiers, i.e. [-D, +Cl_{INTR}] (88iii), and another that has Ds but lacks the lexical rule for classifiers, i.e. [+D, -Cl_{INTR}] (88iv).

(88) Predicting CILs with [\pm D, \pm Cl_{INTR}]

	CILs	D	Cl _{INTR}
(i)	Mandarin	-	-
(ii)	Yi	+	+
(iii)	?	-	+
(iv)	?	+	-

We expect [-D, +Cl_{INTR}] CILs (88iii) to have three main properties. First, their bare nouns should denote kinds and behave similarly to those in Mandarin and the rest of the CILs. Second, their NCs should behave similarly to those in Mandarin and other languages. Note that in CILs *with* Ds, NCs have the option to shift to a definite interpretation via the definite D, as in Yi. However, in CILs *without* D, this option is not available. Third, since this type of CILs possesses the intransitivization rule for classifiers, bare CIPs in these CILs are expected to appear freely in argument position and receive a definite interpretation despite the absence of D (due to Ranking of Meanings in (67)).

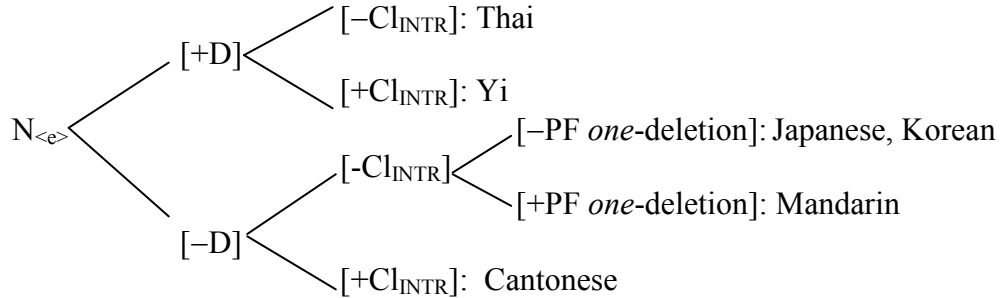
As for [+D, -Cl_{INTR}] languages (88iv), we expect them to have the following properties. First, they have no unrestricted bare CIPs in virtue of lacking the lexical intransitivization rule for classifiers. Second, their bare nouns should denote kinds and behave similarly to those in Mandarin and the rest of CILs. Third, their NCs should behave similarly to those in Mandarin and other languages. Fourth, since the only type of property-denoting nominal is NCs, D should be detected only at the NC level, turning NCs into arguments with a definite interpretation.

We find that Cantonese matches the properties in (88iii) and Thai matches the properties in (88iv) (see also Jiang 2012).³¹ In addition, Mandarin allows *one*-deletion to apply to its NCs in a restricted manner during the process of externalization at PF. In principle, there could be a CIL similar to Mandarin in having the [-D, -Cl_{INTR}] setting but different from it in disallowing *one*-deletion; we find that Korean/Japanese instantiates this type (i.e. its bare CIPs are simply banned (21)). The proposed

³¹ Relevant Cantonese examples can be found in Cheng and Sybesma 1999, 2005, Simpson 2005, and Jiang 2012; relevant Thai examples can be found in Jenks 2011 and Jiang 2012.

comparative analysis of Mandarin and Yi, thus, has rich consequences for the typology of argument formation, as summarized below.

(89) Updated Typology of argument formation



6 Summary and concluding remarks

The main purpose of this paper was to understand how nominal arguments are formed in CILs and whether or not nouns universally require Ds when serving as arguments. We studied comparatively two types of CILs (instantiated by Mandarin and Yi respectively) and contrasted them with NMLs like English, French, and Hindi. We showed that the following two phenomena: i) whether or not a language has overt lexical articles; and ii) whether or not the language allows bare arguments, yield three types of NMLs but only two types of CILs. The first type of CILs, as exemplified by Mandarin, resembles NMLs like Hindi in that it freely allows bare arguments and does not have Ds. The second type of CILs, as exemplified by Yi, is previously unattested and is superficially similar to NMLs like English in that it freely allows bare arguments and also has overt Ds. As we have shown, studying the two types of CILs enables us to understand why there is no analogue of the third type of NMLs like French among CILs (i.e. why it is the case that *no* known CIL requires overt Ds to argumentize bare nouns) and why D develops in languages where it is not required.

We observed that although they differ a great deal in the internal nominal domain, NCs in NMLs and CILs (e.g. in the form of *two boys* and *liang ge nansheng* ‘two Cl boy’) share strong similarities at the clausal level concerning their long-distance scope behavior, semantic interpretation and syntactic distributions. Based on such facts, we argued for a lexical view of numerals that treats them as ambiguous: on one hand, numerals are property modifiers, type $\langle\langle e, t \rangle, \langle e, t \rangle\rangle$; on the other hand, each numeral has a lexically predictable alternant of type $\langle\langle e, t \rangle, e \rangle$, formed via a lexical rule. According to this lexical analysis of numerals, NCs are systematically ambiguous (in the Universal Lexicon) between a predicate and an indefinite variant. In their predicate incarnation, NCs act as restrictors of Ds/demonstratives, quantifiers, and the generic operator. In their indefinite incarnation, NCs are arguments with long-distance scope properties and do *not* project DPs in the syntax.

Language variation in the nominal domain, as we argued, is primarily located in two interrelated factors: what nouns denote and what low functional heads (i.e. number morphology and classifiers) denote. Specifically, we argued that bare nouns in CILs denote kinds, type $\langle e \rangle$, and that classifiers are relations between numerals and atomized noun denotations. In contrast, it was argued that nouns in NMLs denote properties and that number morphology is a function from properties to properties.

Given that bare nouns in CILs are always argumental, we expect that any CIL always allows bare nouns to merge directly with verbs and occur freely in argument position. Furthermore, if Ds are universally property taking (of, say, type $\langle et, e \rangle$), we should not expect them to combine with bare nouns in CILs. Therefore, the proposed analysis correctly predicts that an obligatory CIL in which nouns cannot occur bare as arguments (the analogue of French) should not exist. Nevertheless, if some higher nominal projections above the bare noun, namely bare CIPs and NCs, are of a predicative type ($\langle e, t \rangle$), lexical determiners become relevant and may well develop to turn them into arguments. Hence, the proposed analysis further predicts that if a CIL develops a determiner, this element will be able to apply only at the level of NCs, which are property-denoting, but not at the level of bare nouns, which are kind-denoting, given type theoretic considerations. If an intermediate projection between NCs and bare nouns that is of type $\langle e, t \rangle$, namely a bare CIP, is available, determiners can apply at this level as well.

Importantly, we reduced the micro-variation between Mandarin and Yi in the nominal domain to two main parameters: $[\pm Cl_{INTR}]$ in the lexicon, and $[\pm D]$ in the syntax. Mandarin was argued to be a D-less CIL which lacks the lexical ‘intransitivization’ process (INTR) for classifiers that derive unrestricted bare CIPs; whereas Yi was argued to be a CIL with Ds that also has INTR. Our analysis, hence, provides further evidence for the view that there is variation and parameterization in the obligatory status of D (e.g. Chierchia 1998; Dayal 2004; Bošković 2005 et seq).

We showed that a series of further consequences follow naturally from the proposed analysis of nominal arguments. First, semantically, demonstratives can be obtained either from properties or from kinds via the help of indexically individuated situations and thus, unlike determiners, demonstratives can apply at different levels in nominal structures. Second, the same strategy (minus the indexical component) can be exploited to obtain definite readings for bare nouns. Third, language variation regarding the definite reading of bare nouns in CILs can be explained by some competition mechanisms that apply to bare nouns and bare CIPs. Last but not least, we speculated that our two parameters, i.e. $[\pm Cl_{INTR}]$ and $[\pm \text{lexical } D]$ predict further CIL types that are indeed attested.

The study of more languages with different nominal systems may lead to an essential reorganization of the current picture. For the time being, however, it seems to us

that studying Mandarin and Yi comparatively has helped us to get a few steps closer to a general theory of argument formation of wide cross-linguistic applicability.

Appendix

- (1) Three additional common properties of NCs in French, Russian, Mandarin, and Yi
 - (i) Property 4: Lack of anaphoric uses
 - (ii) Property 5: Long-distance scope interpretation and island-escaping ability
 - (iii) Property 6: Generic construals

- (2) Property 4: Lack of anaphoric use of NCs

a. *Jean a acheté deux chiens et deux chats.* (French)

John has bought two dogs and two cats

#(Les/ces) deux chats sont coûteux.

The/these two cats are expensive

‘John bought two dogs and two cats. The/these two cats are expensive.’

b. *Ivan kupil pjat' sobak i pjat' koshek.* (Russian)

Ivan bought five dogs and five cats.

#(Eti) pjat' koshek ochen' dorigie.

These five cats very expensive

‘Ivan bought five dogs and five cats; these five cats are very expensive.’

c. *fang-jian li zuo zhe san ge nansheng he wu ge nvsheng,*
room inside sit Prog three Cl boy and five Cl girl
#(na) wu ge nvsheng hen chao-si le. (Mandarin)

that/those five Cl girl very noisy-dead SFP

‘Three boys and five girls are sitting in the room. Those five girls are so noisy.’

d. *si-hni sɔ ma sini sse-vo sɔ ma i go nyi,*
girl three Cl and boy three Cl house sit,
si-hni #(a daɿ) sɔ ma / si-hni sɔ ma #(su) dzi ndza. (Yi)

girl that three Cl girl three Cl SU very beautiful

‘Three girls and three boys are sitting in the house; those/the three girls are very pretty.’

- (3) Property 5: Long-distance scope and island-escaping ability of NCs

a. *Si tu apportes deux filles à la fête, j'en serai heureux.* (French)

If you bring two girls to the festival I will happy

‘If you bring two girls to the party I will be happy’

i. Wide scope: [two > if] [island escaping: wide scope]

ii. Narrow scope: [if > two]

- a'. *La plupart des professeurs ont lu chaque essai que deux étudiants ont écrit.*
 the most of professors have read every essay that two students have written
 'Most professors have read every paper that two students wrote.'
 i. Intermediate scope: [most>two>every] [island escaping: intermediate scope]
 ii. Narrow scope: [most>every>two]
- b. *Ja budu schastliv, esli ty pridjosh' na vecherinku s dvumja devushkami.*
 I be.fut happy if you come on party with two girls
 'If you bring two girls to the party, I will be happy.' (Russian)
 i. Wide scope: [two>if] [island-escaping: wide scope]
 ii. Narrow scope: [if>two]
- b'. *Bol'shinstvo professorov prochitali kazhduju rabotu chto napisali dva studenta*
 most professors read every paper that wrote two students
 'Most professors have read every paper that two students wrote.'
 i. Intermediate scope: [most>two>every] [island-escaping: intermediate scope]
 ii. Narrow scope: [most>every>two]
- c. *ruguo ni neng dai yi ge nvsheng lai wo-de party de-hua, (Mandarin)*
 if you can bring one Cl girl come my party if
wo hui hen kaixin.
 I will very happy
 'If you can bring one girl to my party, I will be very happy.'
 i. Wide scope [one girl > if] [island-escaping]
 ii. Narrow scope: [if > one girl]
- c'. *Da-bu-fen laoshi dou jian-le [mei yi ge*
 most teacher DOU meet-Asp every one Cl
[bei san suo changqingteng daxue luqu le De] xuesheng]
 Pass three Cl Ivy League college admit Asp de student
 'Most teachers have met every student who was admitted by three Ivy League schools.'
 i. Intermediate scope: [most >three> every] [island-escaping]
 ii. Narrow scope: [most > every >three]
- d. *ne si-hni nyip zha shyp ngop ddu ggep la go-li, (Yi)*
 you girl two Cl bring my home play come if,
nga khat ox.
 I happy SFP
 'If you bring two girls to come to hang out in my home, I will be happy.'
 i. Wide scope [two girls > if] [island-escaping]
 ii. Narrow scope: [if > two girls]

(4) Property 6: Generic/individual level interpretation of NCs in NMLs

- a. *Deux garçons peuvent soulever un piano.* (French)

two boys can raise a piano
 ‘Two boys can lift a piano.’

b. *Pjat' parnej ne mogut podnjat' pianino.* (Russian)
 five boys NEG can lift piano
 ‘Five boys cannot lift a piano.’

c. *san ge nvsheng keyi tai-qi yi jia gangqin* (Mandarin)
 three Cl girl can lift-up one Cl piano
 ‘Three girls can lift a piano.’

d. *mu so ma so tu va bi.* (Yi)
 horse three Cl three thousand dollar give
 ‘Three horses cost three thousand dollars.’

References

- Abusch, Dorit. 1994. The Scope of Indefinites, *Natural Language Semantics* 3, 88-135.
- Bale, Alan, Michael Gagnon, and Hrayr Khanjian. 2011. Cross-linguistic representations of numerals and number marking. *Proceedings of SALT 20*: 582-598.
- Barwise, J. & Cooper, R. (1981) ‘Generalized quantifiers and natural language’. *Linguistics and Philosophy* 4:159–219.
- Bhattacharya, Tanmoy 1999a. *Structure of the Bangla DP*. University College, London. PhD dissertation.
- Bhattacharya, Tanmoy 1999b. DP internal specificity in Bangla. *Yearbook of South Asian Languages and Linguistics* 2. 71-99.
- Borer, Hagit. 2005. *Structuring sense, volume 1: In name only*. Oxford: Oxford University Press.
- Bošković, Željko. 2005. On the locality of left branch extraction and the structure of NP. *Studia Linguistica* 59: 1-45.
- Bošković, Željko. 2007. On the locality and motivation of move and agree: An even more minimal theory. *Linguistic Inquiry* 38:589-644.
- Bošković, Željko. 2008. What will you have, DP or NP? In *Proceedings of 37th Conference of the North-Eastern Linguistic Society*, 101-114. Amherst: GLSA, University of Massachusetts.
- Carlson, Greg N. 1977a. A unified analysis of the English bare plurals. *Linguistics and Philosophy* 1: 413-456.
- Carlson, Greg N. 1977b. Reference to kinds in English. Doctoral Dissertation, University of Massachusetts, Amherst.
- Chao, Yuen Ren. 1968. *A grammar of spoken Chinese*. Berkeley: University of California Press.
- Cheng, Lai-Shen, Lisa & Rint, Sybesma. 1999. “Bare and not-so-bare nouns and the structure of NP,” *Linguistic Inquiry* 30(4): 509-542.
- Cheng, Lai-Shen, Lisa & Rint, Sybesma. 2005. Classifiers in four varieties of Chinese." In Guglielmo Cinque and Richard Kayne, *The Oxford handbook of comparative*

- syntax*, 259-292. New York: OUP USA.
- Chierchia, Gennaro. 1992. "Anaphora and dynamic logic." *Linguistics and Philosophy* 15:111-183.
- Chierchia, Gennaro. 1995. Individual-level predicates as inherent generics. In *The generic book*, eds. Gregory N. Carlson and Francis Pelletier, 176-223. Chicago: University of Chicago Press.
- Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6:339-405.
- Chomsky, Noam. 1994. *Bare phrase structure*. Cambridge, MA: MITWPL.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2000 Minimalist inquiries. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89-155. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1-52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2004. Beyond explanatory adequacy. In *Structures and beyond*, ed. Adriana Belletti, 104-131. Oxford: Oxford University Press.
- Dahl, Östen. 1975. On generics. In E. Keenan, ed., *Formal Semantics of Natural Language*, 99-111. Cambridge: Cambridge University Press.
- Dayal, Veneeta. 2004. Number Marking and (in)definiteness in kind terms. *Linguistics and Philosophy* 27:4, 393-450.
- Dayal, Veneeta. 2010. Bangla Classifiers: Mediating between kinds and objects. presented at the Conference on Bantu, Chinese & Romance DPs. University of Leiden.
- Dayal, Veneeta. 2011a. Bare Noun Phrases. Ms. Rutgers University. To appear in Maeinborn, von Heusinger and Portner (eds), *Semantics: An International Handbook of Natural Language Meaning*, 33. 2 1087-1108.
- Dayal, Veneeta. 2011b. Bangla classifiers: mediating between kinds and objects. Paper presented at *International Workshop on Syntax-Semantics Interface*. Institute of Linguistics, Academia Sinica, Taiwan.
- Dayal, Veneeta. To appear. Bangla classifiers: mediating between kinds and objects. *Italian Journal of Linguistics*.
- Declerck, Renaat. 1986. The manifold interpretations of generic sentences. *Lingua*, 68, pp.149-188.
- Di Sciullo, Anna Maria. 2012. Asymmetry in complex numerals. Talk given at Harvard University.
- Doetjes, Jenny. 1997. Quantifiers and selection: On the distribution of quantifying expressions in French, Dutch and English. Doctoral dissertation, Leiden University.
- Frank, Steven. 1994. Parametric properties of numeral phrases in Slavic. *Natural Language & Linguistic Theory* 12:597-674.
- Farkas, Donka F. 1981. Quantifier scope and syntactic islands. *Chicago Linguistics Society* 17. 59-66.

- Fodor, J. D. & Sag, I. (1982) 'Referential and quantificational indefinites'. *Linguistics and Philosophy* 5:355–398.
- Giusti, Giuliana. 1991. The categorial status of quantified nominals. *Linguistische Berichte* 136:438–452.
- Giusti, Giuliana. 1997. The categorial status of determiners. In L. Haegeman (ed), *The New Comparative Syntax*. Cambridge University Press. Cambridge. 94–113.
- Haegeman, Liliane & Guéron, Jacqueline. 1999. *English grammar: A Generative Perspective*. Blackwell. Oxford.
- Heim, Irene. 1982. The semantics of definite and indefinite noun phrases. Doctoral Dissertation. University of Massachusetts.
- Heim, Irene. 1989. Articles and Definiteness. Ms. UCLA. (Published in German as "Artikel und Definitheit." (1991) *Semantics: An International Handbook of Contemporary Research*, eds. Arnim von Stechow and Dieter Wunderlich, 487–535. Berlin: de Gruyter, 2.)
- Huang, C.-T. James. 1997. On lexical structure and syntactic projection. *Chinese Languages and Linguistics* 30: 45-89.
- Huang C.-T. James. 2009. Variation and change in parametric theory: an East Asian perspective. Harvard-MIT handout.
- Ionin, Tania, and Ora Matushansky. 2006. The composition of complex cardinals. *Journal of Semantics* 23, 315-360.
- Jenks, Peter. 2011. The hidden structure of Thai noun phrases. Doctoral Dissertation, Harvard University.
- Jiang, L. Julie and Suhua Hu. 2010 An overt determiner in a classifier language, presented at GLOW-in-Asia VIII, Beijing, China. (to appear in *Proceedings of GLOW-in-Asia VIII*)
- Jiang, L. Julie. 2012. Nominal arguments and language variation. Doctoral Dissertation, Harvard University.
- Kennedy, Christopher. 1999. Projecting the adjective: The syntax and semantics of gradability and comparison. New York: Garland.
- Kennedy, Christopher, and Louise McNally. 2005. Scale structure, degree modification, and the semantics of gradable predicates. *Language* 81: 345–381.
- Koopman, Hilda, and Dominique Sportiche. 1991. The Position of Subjects. *Lingua* 85. 2/3.
- Krifka, Manfred. 1995. Common nouns: A contrastive analysis of Chinese and English. In *The generic book*, ed. Gregory Carlson and Francis Pelletier, 398–411. Chicago: University of Chicago Press.
- Krifka, Manfred. 1999. At least some determiners aren't determiners. In K. Turner (ed.), *The semantics/pragmatics interface from different points of view*. (= Current Research in the Semantics/Pragmatics Interface Vol. 1). Elsevier Science B.V., 257-291.
- Krifka, Manfred. Francis Jeffry Pelletier. Gregory N. Carlson, Alice ter Meulen, Godehard Link, and Gennaro Chierchia. 1995. Genericity: An introduction. IN: Gregory N. Carlson. And Francis J. Pelletier. (eds.). *The generic book*. Chicago: University of Chicago Press.
- Kratzer, Angelika. 1981. The notional category of modality. In *Words, worlds, and*

- contexts, eds. H. J. Eikmeyer and H. Rieser, 38-74. New York: de Gruyter.
- Kratzer, Angelika. 1998. Scope or pseudoscope? Are there wide scope indefinites? In S. Rothstein (ed), *Events and Grammar*. Kluwer. Dordrecht. 163–196.
- Landman, Fred. 2003. Predicate-argument mismatches and the adjectival theory of indefinites. In M. Coene & Y. D’hulst (eds), *The Syntax and Semantics of Noun Phrases. Linguistics Today 55*. John Benjamins. Amsterdam and Philadelphia. 211–237.
- Lawler, John. 1972. Generic to a fault. *Chicago Linguistics Society 9*, 320-331.
- Li, Charles, N. & Sandra, A. Thompson. 1976. Subject and Topic: A New Typology of Language. IN: Charles N. Li. (eds). *Subject and Topic*. New York: Academic Press. 458-489.
- Li, Xu-Ping. 2011. On the Semantics of Classifiers in Chinese. Doctoral Dissertation, Bar Ilan University & University of Mainz.
- Li, Xuping and Walter Bisang. 2012. Classifiers in Sinitic languages: from individuation to definiteness-marking. *Lingua 22*: 335-355.
- Li, Yen-hui Audrey. 1997. Structures and Interpretations of Nominal Expressions. Ms. University of Southern California, LA, California.
- Li, Yen-hui Audrey. 1998. Argument Determiner Phrases and Number Phrases. *Linguistic Inquiry*, Vol. 29, No. 4, pp.693-702.
- Li, Yen-hui Audrey. 1999. Plurality in a classifier language. *Journal of East Asian Linguistics 8*:75–99.
- Li, Yen-hui Audrey and Shengli Feng. 2013. Beijinghua ‘yi’ zi shenglüe de yunlü tiaojian [the licensing conditions for the deletion of ‘one’ in Mandarin]. Ms. University of Southern California and Chinese University of Hong Kong.
- Link, Godhard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach, in R. Bauerle, C. Schwartze and A. von Stechow (eds.), *Meaning, Use and Interpretation of Language*, 302-323, de Gruyter.
- Link, Godhard. 1987. Generalized quantifiers and plurals. In P. Gärdenfors (ed), *Generalized Quantifiers*. D. Reidel. Dordrecht. 151–180.
- Longobardi, Giuseppe. 1994. Reference and proper names: a theory of N-movement in syntax and logical form. *Linguistic Inquiry 25*:pp.609-665.
- Longobardi, Giuseppe. 2001: How comparative is semantics?, *Natural Language Semantics 9*: 335-369.
- Lü , Shuxiang. 1944. ‘Ge’ zi de yingyong fanwei, fulun danweici qian ‘yi’ zi de tuoluo. (On the usages of GE and the dropping of the numeral yi before classifiers). *Xihua Daxue zhongguo wenhua huikan*, Jinlin, Qilu.
- Partee, Babara. H. 1986. Noun phrase interpretation and type-shifting principles. In J. Groenendijk, D. de Jongh & M. Stokhof (eds.), *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*. GRASS 8. Foris. Dordrecht. 115–143.
- Partee, Barbara. 1987. Noun phrase interpretation and type-shifting principles. In J. Groenendijk, D. de Jongh & M. Stokhof (eds.), *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*. GRASS 8. Foris. Dordrecht. 115–143.
- Progovic, Ljiljana. 1998. Determiner phrase in a language without determiners. *Journal*

- of Linguistics* 34: 165-179.
- Reinhart, Tanya. 1997. Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy* 20:335-397.
- Ruys, Eddy. G. 1992. *The Scope of Indefinites*. Unpublished PhD thesis, Utrecht University.
- Sauerland, Uli. 2003. A New Semantics for Number. In R: Young and Y: Zhou (eds.) *Proceedings of SALT* 13.
- Sauerland, Uli, Jan. Anderssen, and Kazuko Yatsushiro. 2005. The plural is semantically unmarked. In S: Kesper and M: Reis (eds.) *Linguistic Evidence*. Berlin: Mouton de Gruyter.
- Scontras, Greg. 2011. Measure Phrases and Number Marking. Ms. Harvard University.
- Selkirk, Elisabeth. 1977. Some remarks on noun phrase structure. In P. W. Culicover, T. Wasow & A. Akmajian (eds), *Formal Syntax*. Academic Press. London. 285–316.
- Simpson, Andrew. 2005. Classifiers and DP structure in southeast Asian languages. In *The oxford handbook of comparative syntax*, eds. Guglielmo Cinque and Richard Kayne, 806-838. Oxford: Oxford University Press.
- Simpson, Andrew, Hooi Ling Soh, Giang Le, and Hiroki Nomoto. 2011. Bare classifiers and definiteness: a cross-linguistic investigation. *Studies in Language* 35: 1, 168-193.
- Ritter, Elizabeth. 1991. Two functional categories in noun phrases: Evidence from Modern Hebrew. *Perspectives on Phrase Structure. Syntax and Semantics* 25. Academic Press. New York. 37–62.
- T'sou, Benjamin K. 1976. The structure of nominal classifier systems. In *Austroasiatic Studies*, vol 2, (eds.) Philip N. Jenner, Stanley Starosta, and Laurence C. Thompson. Honolulu: University Press of Hawaii, 1215-47.
- Tang, Chih-Chen. Jane. 1990. A note on the DP analysis of the Chinese noun phrase. *Linguistics* 28:337-354.
- Trinh, Tue. 2010. Nominal reference in two classifier languages. In *Sinn und Bedeutung* 15.
- Tsai, Cheng-Yu. Edwin . 2011. Some problems on NumP and more problems on DP. Ms. Harvard University.
- von Fintel, Kai. 1994. Restrictions on Quantifier Domains. Doctoral Dissertation, University of Massachusetts, Amherst.
- Vergnaud J.-R. & M.L. Zubizarreta. 1992. The definite determiner and the inalienable constructions in French and in English. *Linguistic Inquiry* 23, 595-652.
- Winter, Yoad. 1997, Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy* 20:399–467.
- Wu, Yicheng, and Adams Bodo. 2009. Classifiers ≠ determiners. *Linguistic Inquiry* 40:487–503.
- Yang, Rong. 2001. *Common nouns, classifiers, and quantification in Chinese*. Doctoral dissertation, Rutgers University.
- Zamparelli, Roberto. 1995. *Layers in the Determiner Phrase*. Unpublished PhD thesis, University of Rochester.
- Zabbal, Youzi. 2005. The syntax of numeral expressions. Ms, University of Massachusetts, Amherst.
- Zeijlstra, Hedde. 2004, *Sentential Negation and Negative Concord*, PhD Dissertation,

University of Amsterdam.
Zhang, Niina Ning. 2013. *Classifier Structures in Mandarin Chinese*. Berlin & New York:
Mouton de Gruyter.