Articulated definiteness without articles

Peter Jenks

Abstract

While it lacks a definite article, Mandarin makes a principled distinction between unique and anaphoric definites: unique definites are realized with a bare noun, anaphoric definites are realized with a demonstrative, except in subject position. The following proposals account for these facts: 1) bare nouns achieve definite interpretations via a last-resort type-shifting operator ι, which has a unique definite meaning, 2) demonstratives can occur as anaphoric definiteness because they have a semantic argument beyond their nominal restriction that can be filled by an index, and 3) bare nominal subjects are topics. A principle called *Index!* requires that indexical expressions be used whenever possible. Mandarin is contrasted with Cantonese, which, like English, is shown to have access to a single, ambiguous, definite article.[1]

1 Introduction

Many of the world's languages lack definite articles. Instead, bare nouns or other nominal expressions are found in definite contexts. Because articles are taken to head a DP projection in languages that have them (Abney 1987; Szabolcsi 1994), a relatively obvious analysis of languages without definite articles is that they lack a DP projection altogether (Boškoviç 2008, 2012). Such 'NP-analyses' have been proposed to be generally applicable to numeral classifier languages like Mandarin (e.g. Chierchia 1998; Boškoviç and Hsieh 2012; Cheng 2013) and Japanese (e.g. Fukui and Takano 2000), languages where bare nouns can receive definite interpretations. But if there is no DP, an alternative source for definite interpretations must be proposed. One option has been a semantically supplied definite type-shifting operation, proposed to be universally available as a kind of repair (Chierchia 1998; Dayal 2004, 2011). Another suggestion is that that numeral classifiers themselves serve some of the functions of definite articles, and hence might be the correlate of definite articles in classifier languages (Cheng and Sybesma 1999, 2012). Still

---

others have maintained that the mapping between DP syntax and definite semantics is universal, concluding by hypothesis that definite bare nouns always involve a null determiner (Simpson 2005; Wu and Bodomo 2009).

At the same time, the empirical picture which these theories must cover has become increasingly complex. For one, while most numeral classifier languages lack articles, they express definiteness differently. In their landmark paper on noun phrases and definiteness in Chinese, Cheng and Sybesma (1999) observe that while Mandarin uses bare nouns in subject and object position to achieve definite interpretations, Cantonese must include a numeral classifier in equivalent environments:

(1)　MANDARIN: DEFINITE BARE NOUNS (Cheng and Sybesma 1999:510)

    a.　*Hufei he-wan-le　　tang.*
       Hufei drink-finish-PERF soup
       'Hufei finished the soup.'

    b.　*Gou yao guo malu.*
       dog want cross road
       'The dog(s) want to cross the road.'

(2)　CANTONESE: DEFINITE [CLF+N] PHRASES (Cheng and Sybesma 1999:511)

    a.　*Zek gau zungji sek juk*
       CLF dog like　eat meat
       'The dog likes to eat meat.'

    b.　*Ngo zungji tong zek gau waan.*
       I　　like　with CLF dog play
       'I like to play with the dog.'

The distinction in definiteness marking between Mandarin and Cantonese is echoed in a very general typological split among numeral classifier languages. While many languages such as Japanese (Kurafuji 2004), Korean (Kim 2005), and Thai (Piriyawiboon 2010) pattern more or less with Mandarin, others, including Vietnamese (Löbel 2000; Nguyen 2004; Trinh 2011) Hmong (Bisang 1999), and Bangla (Dayal 2012; Simpson and Biswas 2016) pattern more or less with Cantonese. In still other numeral classifier languages, definiteness is marked overtly, either inflectionally on the classifier, as in such as Weining Ahmao (Gerner and Bisang 2010) and Wenzhou Wu Chinese,(Cheng and Sybesma 2005; Sio 2006), or with a separate definite article, as in Nuosu (or Yi)(Jiang 2012).

In this literature, the notion of definiteness is typically treated as a semantic primitive, and 'definite environments' describe the general distribution, for example of [Clf-N] phases in Cantonese, Mandarin bare nouns, and definite descriptions in English. The main empirical contribution of this paper is to demonstrate that by paying more attention to different types of definite environments we find that the distribution of Mandarin bare nouns is different from English definite descriptions and Cantonese [Clf-N] sequences, and that Mandarin demonstratives play a central role in the expression of definiteness. More specifically, bare nouns in Mandarin are only available in definite environments licensed by uniqueness,

while demonstratives occur in most anaphoric environments. This is illustrated in the following Mandarin donkey sentences with an anaphoric full noun phrase, where a bare noun is impossible, but a demonstrative is fine:

(3) a. MANDARIN DONKEY SENTENCE WITH BARE NOUN

 mei ge [you yi zhi shuiniu de] nongfu dou hui da <u>shuiniu</u>.
 every CLF have one CLF buffalo REL farmer all   will hit buffalo

 'Every farmer that has a buffalo hits buffalo (generally).' (no bound reading)

  b. MANDARIN DONKEY SENTENCE WITH DEMONSTRATIVE

 mei ge [you yi zhi shuiniu de] nongfu dou hui da <u>na zhi   shuiniu</u>.
 every CLF have one CLF buffalo REL farmer all   will hit that CLF buffalo

 'Every farmer that has [a buffalo]$_i$ hits [that buffalo]$_i$.'

In anaphoric environments like this one, on the other hand, Cantonese allows [Clf-N] phrases (4), just as English allows a definite article (Elbourne 2013) (5):

(4) CANTONESE DONKEY SENTENCE WITH [CLF+N] PHRASE

 mui go jau jat zek maa ge lungfu daa <u>zek</u> <u>maa</u>.
 every CLF have one CLF horse REL farmer hit CLF horse

 'Every farmer that has [a horse]$_i$ hits [that horse]$_i$.'

(5) Every farmer that has a donkey beats the donkey.

We will see that Mandarin systematically uses bare nouns in unique positions and demonstratives in anaphoric positions with one exception: subject position, where bare nouns seem to be available in Mandarin even when anaphoric.

I offer an analysis of these facts which is consistent with the idea that Mandarin lacks a definite article altogether. Instead, Mandarin makes use of type-shifting to achieve definite interpretations of bare nouns (Chierchia 1998; Yang 2001; Dayal 2004, 2011). The semantics of this type-shift must only be compatible with unique definite environments, however. I suggest this is because type-shifting cannot introduce additional arguments such as an index. As such, demonstratives are possible in these contexts because they are capable of introducing such indices. Because unique definites should be in principle supply the right individual even in anaphoric environments, I suggest that a principle called *Index!* requires that an indexical expression be used when possible.

Turning back to Cantonese, I show towards the end of the paper that [Clf-N] phrase in Cantonese occur in both unique and anaphoric definite environments, as does English *the*, and conclude from this that both languages have a definite article which is ambiguous between unique and anaphoric semantics. I then survey the typology of definiteness marking, and show that it is quite generally true that when definite bare nouns are semantically restricted it is to unique definite environments. This correlation suggests that indices are only available when a DP is syntactically projected across languages.

3

The outline of this paper is as follows: Sections 2 provides background discussion of definiteness. Sections 3 lays out the distribution of bare nouns and demonstrative descriptions in Mandarin. Section 4 offers a syntactic and semantic analysis of these expressions, including empirical evidence for an overt index argument in Mandarin. Section 5 discusses the competition between these two forms, and the various grammatical mechanisms which regulate their competition including when bare nouns are topics. Section 6 shows that Cantonese [Clf-N] phrases, like English *the*, is used generally in definite environments, and that this provides further support for the proposed analysis of Mandarin as well as informing a typology of definiteness. Section 7 concludes.

2 Varieties of definiteness

A central debate in the literature on definiteness is whether it is better characterized by uniqueness or familiarity.[2] Uniqueness theorists (e.g. Russell 1905; Kadmon 1990; Hawkins 1991) can easily account for the use of definite articles in noun phrases such as *the sun*, *the Prime Minister of England* or *the period at the end of the sentence*, all of which are unique in some relevant context. Additionally, the definite article in these contexts seems to be licensed by uniqueness alone; no prior mention of suns or Prime Ministers is needed.

Because they require no prior mention in discourse, uses of the definite article licensed by uniqueness pose problems for familiarity-based analyses. Familiarity-based views of definiteness have been a central component in dynamic theories of meaning (Kamp 1981; Heim 1982; Kamp and Reyle 1993; Chierchia 1995). These theories assume that definite descriptions are variables interpreted relative to a contextually supplied assignment function. Narrative sequences like (6) provide evidence for the role of familiarity:

(6)   The tycoon complained to the senators that a gnome vandalized his resort, and alleged that the gnome used a flamethrower.

The definite article in the second sentence is licensed by the indefinite in the first clause. Crucially, there might not actually be a unique gnome in the context where this sentence is uttered; it could be a well-documented fact that the resort was awash in gnomes.[3]

Uniqueness theorists and familiarity theorists have argued that their respective notions of definiteness can extend to the paradigm cases for the alternative perspective, or have blended the two theories to the point where they are no longer clearly distinguishable. For example, Roberts (2003) offers a hybrid view of definiteness, arguing that cases like *the sun* meet a criterion she calls *weak familiarity*, which is a presupposition on the use of definites. At the same time, she argues that the definite article in cases like (6) satisfy uniqueness in the limited scope of the preceding context, thus preserving both components of the tradi-

---

[2] See Schwarz (2009, ch. 1) and Abbott (2010, 214-226) for recent overviews of this debate.

[3] Lyons (1999) argues that familiarity is a subtype of a larger class of definiteness licensed by identifiability; because the notion of identifiability per se has not fully incorporated into formal semantic treatments of definiteness, I will set this important observation aside here.

tional view. Extending aspects of this insight, although primarily in a uniqueness-based view, Schwarz (2009) and Elbourne (2013) argue that uniqueness in definite noun phrases must be relativized to specific contexts of interpretation, *minimal situations* in the sense of Kratzer (1989, 2007).

Yet recent findings in languages besides English have challenged the conventional view that there a single uniform analysis of definiteness for all of language is desirable, or even possible. For example, Schwarz (2009) describes a distinction between 'strong' vs. 'weak' definite articles in German which tracks familiar versus unique definite environments. The morphological contrast that Schwarz discusses is subtle, only detectable in whether definite articles can contract with prepositions. Weak definite articles, which occur in unique definite contexts, must contract, while strong definite articles, which occur in familiar definite contexts, cannot be contracted:

(7)　　WEAK VERSUS STRONG ARTICLES IN GERMAN (Schwarz (2009, 41))

　　　a.　*In der Kabinettssitzung heute wird ein neuer Vorschlag <u>vom</u>　　<u>Kanzler</u>*
　　　　　In the cabinet meeting　today is　a　new　proposal　by-the_{weak} chancellor
　　　　　erwartet.
　　　　　expected
　　　　　'In today's cabinet meeting, a new proposal by the chancellor is expected.'

　　　b.　*In der Kabinettssitzung heute wird ein neuer Vorschlag #<u>vo-m</u>/*
　　　　　In the cabinet meeting　today is　a　new　proposal　by-the_{weak}/
　　　　　<u>*von dem*</u>　　<u>*Minister*</u> *erwartet.*
　　　　　by　the_{strong} minister expected
　　　　　'In today's cabinet meeting, a new proposal by the minister is expected.'

In (7a), because chancellors are unique, the definite article is weak and must be contracted with the preposition. But as there are can be many ministers of parliament, say in Germany, contracting the definite article before 'minister' is in (7b) results in infelicity and an anaphoric meaning is triggered. If the contracted preposition in (7a) is replaced with the full pronoun and article (*von dem*), the sentence is acceptable, but only in a context where we have mentioned some chancellor. So the strong article is always anaphoric to an existing discourse referent.

Arkoh and Matthewson (2013) describe a similar distinction in a West African language, Fante (Akan). The Fante definite article *nʊ*, which was characterized as optional in previous literature, occurs with familiar noun phrases but not in contexts where definiteness is licensed only by uniqueness:

(8)　　FANTE FAMILIAR DEFINITE ARTICLE (Arkoh and Matthewson 2013, 34,2)

　　　a.　*mʊ-rʊ-kɔ*　　*gua*　*mu*
　　　　　1SG.SUBJ-PROG-go market in
　　　　　'I am going to the market.'

　　　b.　*mʊ-tɔ-kɔ*　　*ekutu.　Ekutu* **(nʊ)*　*yɛ dɛw papa*
　　　　　1SG.SUBJ-buy-PAST orange. Orange *(FAM) be nice good

'I bought an orange. The orange was really tasty.'

The object of (8a) is a uniqueness definite, and the definite article is absent. In contrast, the anaphoric subject of the second clause in (8b) requires the definite article, which follows the noun. Thus, bare nouns in Fante are unique definites, like German weak articles, while the Fante definite article is licensed by familiarity, like German strong articles.

It turns out that morphosyntactic distinctions between uniqueness and familiarity are quite common. Schwarz (2013) identifies Fering (a dialect of Frisian), Lakhota, and Hausa as languages which make a morphological distinction between weak and strong articles, and Ingason (2016) establishes the same contrast in Icelandic free definite articles.[4] Schwarz (2013) also observes that some languages only mark anaphoric definite environments, including Mauritian Creole and Akan as languages where the definite article is restricted to anaphoric environments. The following section shows that unique versus familiar definiteness is also distinguished in Mandarin: while unique definites must be realized as bare nouns, familiar definites occur with demonstratives.[5]

3   Definiteness in Mandarin

Bare nouns in Mandarin can be definite, meaning that they can be used in contexts where a definite article would be obligatory in English, as in the following example from Cheng and Sybesma (1999):

(9)   *Gou* yao guo   malu.
        dog want cross road
        'The dog(s) want to cross the road.'   (Cheng and Sybesma 1999:510)

At the same time, authors such as Chen (2004) claim that demonstratives in Mandarin can mark definiteness, although it is unclear under what circumstances definite interpretations of demonstratives occur.

This section shows that bare nouns with definite interpretations are restricted to unique definite environments in Mandarin. In contrast, demonstrative descriptions only occur in familiar or anaphoric definite environments, requiring explicitly mentioned discourse antecedents. The obvious conclusion is that Mandarin clearly distinguishes definites licensed

---

[4]Barlew (2014) describes a definite article in Bulu (Bantu, Cameroon) whose distribution is broadly anaphoric, although he argues that this article must refer to salient antecedents, a stronger requirement than mere familiarity associated with the distribution of pronouns (Roberts 2004). This raises the possibility that there is more than one kind of anaphoric definite article across languages.

[5]The distinction above between definites licensed by uniqueness and definites licensed by familiarity is one of many possible distinctions in definite descriptions which must be controlled for. Löbner (1985, 2011) observes, for example, that relational nouns such as *mother* and *weight* are relations from individuals to other unique individuals. This distinction is important and will be controlled for below by only using non-relational common nouns, except in cases of bridging where relational nouns are necessary to produce certain bridged definites.

by uniqueness from those licensed by familiarity, meaning that Mandarin is like the languages surveyed in Schwarz (2013) in distinguishing two kinds of definites.

3.1 Mandarin bare nouns as uniqueness definites

Mandarin bare nouns occur in three environments which are also observed by Schwarz (2009) to require weak definite articles in German: larger situation definites, immediate situation definites, and part-whole bridging.[6] These environments are uniform because definiteness is licensed in them by the pragmatic context and not by prior mention. Demonstratives can only occur in these contexts with contrastive interpretations.

The first definite environment we will examine is larger situation definites, a term due to Hawkins (1978). With larger situation definites, uniqueness is licensed not by the specific conversational context but by general world knowledge. Hence, nouns meaning 'moon' or 'president' are obligatorily definite because there is only one individual for which their descriptive content holds. Larger situation definites in Mandarin are expressed by a bare noun:

(11)   MANDARIN: LARGER SITUATION DEFINITES

    a.   *Yueliang sheng shang lai le.*
           moon   rise   up   come PERF
           'The moon has risen.'   (Chen 2004, p. 1165)

    b.   *(#Na/#zhe ge) Taiwan (de) zongtong hen shengqi*
           that/this   CLF Taiwan MOD president very angry
           'The president of Taiwan is very angry.'

Example (11b) shows that demonstrative determiners are infelicitous in environments licensed only by contextual uniqueness.

Bare nouns in Mandarin also occur as immediate situation definites, a label also due to Hawkins (1978). Immediate situation definites are uniqueness definites that rely on context-

---

[6]Weak definites, on which see Barker (2005) and Carlson et al. (2006), pattern with uniqueness definites in German (Schwarz 2009). Weak definites allow bare nouns in Mandarin, like other unique definite environments:

(10)   MANDARIN: WEAK DEFINITES

    a.   *Zhangsan dai Xiaoli qu yi yuan*
           Z.   take X.   go hospital
           'Zhangsan took Xiaoli to the hospital.'

    b.   *#Zhangsan dai Xiaoli qu jianzhuwu*
            Z.   take Xiaoli go building
            # 'Zhangsan took Xiaoli to the building.'

Yet the notion of a weak definite is not a coherent one in a language like Mandarin where bare nouns allow both definite and scopeless indefinite readings (Cheng and Sybesma 1999; Yang 2001) In other words, the problem of weak definites being non-referential and non-unique is only a problem in languages that have actual definite articles occurring in these contexts.

specific knowledge that is shared by the speaker and hearer. This category has played an especially important role in pragmatic theories of definiteness, including those which rely on notions such as identifiability (e.g. Lyons 1999) and salience (e.g. von Heusinger 2013), because these theories highlight the role of context. Immediate situation definites are common in the existing literature on Mandarin. For instance, Cheng and Sybesma (1999) provide the following examples:

(12)   MANDARIN: IMMEDIATE SITUATION DEFINITES

    a.   *Hufei he-wan-le   tang.*
        Hufei drink-finish-PERF soup
        'Hufei finished the soup.'

    b.   *Gou yao guo malu.*
        dog want cross road
        'The dog(s) want to cross the road.'   (Cheng and Sybesma 1999:510)

These examples evoke a specific utterance context. For example, by virtue of the proper name *Hufei* and aspectual morphology, (12a) must be interpreted in the context of a specific person finishing a specific meal. Similarly, (12b) would not ordinarily be interpreted as a generic statement about dogs, because dogs do not regularly want to cross roads. Instead, this sentence would be expected in the context of a specific dog with a specific intention. The fact that a bare nouns occur as immediate situation definites provides further evidence that bare nouns are uniqueness definites.

Demonstratives cannot be used to express immediate situation definites. If a demonstrative was used in either of the sentences above without prior mention of the noun, it would be accompanied either by a pointing gesture or would be used for contrast with an alternative bowl of soup or dog for whom the predicate was not true.

The contrast between demonstratives and bare nouns is also apparent in bridging definites (Clark 1975) also called associative anaphora (Hawkins 1978) or inferrables (Prince 1981). Clark (1975) distinguishes two instances of bridging or 'indirect reference,' one of 'indirect reference by association' (13a) and another class of 'indirect reference by characterization' (13b):

(13)   a.   I looked into the room. The ceiling was very high.
    b.   John was murdered yesterday. The murderer got away. (Clark 1975, p. 171)

Clark's distinction between association and characterization resembles the split between *part-whole* bridging and *producer-product* proposed by Schwarz (2009, ch. 2):

(14)   a.   Part-whole relationship: room-ceiling, house-living room, etc.
    b.   Producer-product relationship: author-play, painter-painting, etc. (Schwarz 2009, p. 54)

While German judgments are not crystal clear, Schwarz experimentally confirmed that these two kinds of bridging definites occur with different classes of articles in German:

part-whole relationships prefer the weak article, indicating that they are unique definites, while the producer-product associations prefer the strong article, meaning they are familiar definites.

Strikingly, Mandarin bridging contexts track this distinction like German. While definites licensed by part-whole bridging are realized as bare nouns (15a), native speakers I have discussed these data with clearly prefer a demonstrative description in cases of producer-product bridging (15b), meaning they are familiar definites (see below):

(15)   MANDARIN: PART-WHOLE VS. PRODUCER-PRODUCT BRIDGING

    a. *Chezi bei   jingcha lanjie le yinwei mei you tiezhi zai*
        Car   ADV.PAS police intercept PRF because NEG have sticker at
        <u>paizhao</u>   shang
        license plate on
        'The car was intercepted by the police because there wasn't a sticker on the license plate.'

    b. *Paul renwei na shou shi hen youmei, jishi   ta bu   renshi*
        Paul think   that CLF poem very beautiful although he NEG know
        #(<u>na wei)</u> <u>shiren</u>
          that CLF poet
        'Paul thinks that poem is very beautiful although he doesn't know of the poet.'

Part-whole bridging introduces a uniqueness presupposition because the antecedent of the bridged definite entails its existence by virtue of a containment relationship. Using the example above, once the existence of a specific car is established, under normal circumstances, we can assume the existence of a unique license plate. Schwarz (2009) notes that no such containment relationship holds in the case of producer-product bridging. Poems do not contain their poets. Neither do situations containing a poem, such as a poetry reading, entail the existence of a unique poet. Schwarz shows that producer-product bridging in German must have a relational noun as the bridged definite, where the concealed argument of that noun supplies the anaphoric link to the antecedent, an analysis I adopt for Mandarin in Section 4.4. Summarizing, part-whole bridging must rely on pragmatics to satisfy uniqueness, like all unique definites, while producer-product bridging and other cases of indirect reference by characterization rely on an overt anaphoric link.[7]

---

[7]An anonymous reviewer points out that kinship terms allow bridging but do not fall into either category proposed by Schwarz. Kinship terms are almost always bare in Mandarin — even overt possessors are optional. So while I have found that kinship terms are also bare in bridging contexts, it is not clear what this tells us.

3.2 Anaphoric definites in Mandarin

Familiar or anaphoric definites are definites that are anaphoric to an explicit linguistic antecedent.[8] While uniqueness definites must be realized as bare nouns in Mandarin, this section establishes that with the exception of matrix subjects, anaphoric noun phrases must include a demonstrative determiner. Thus, demonstrative determiners in Mandarin will be shown to have a comparable distribution to the strong article in German (Schwarz 2009) and the determiner *nʊ* in Fante (Arkoh and Matthewson 2013).

The simplest type of anaphoric definite are those which occur in narrative sequences (Karttunen 1969, 1976). In these examples, the first sentence introduces a novel discourse referent with an indefinite, and the second example must refer back to this referent with a definite noun phrase. All of our examples will further include two different noun phrases in the first sentence to preclude the use of a pronoun. The following examples illustrate that demonstrative determiners[9] must be used for anaphoric definites in non-subject positions, bare nouns and demonstratives are possible in subject position:[10]

(16)　MANDARIN NARRATIVE SEQUENCES

　　a.　*jiaoshi　li　　zuo-zhe yi　ge　nansheng he　yi　ge　nüsheng,*
　　　　classroom inside sit-PROG one CLF boy　　　and one CLF girl,
　　　　'There is a boy and a girl sitting in the classroom . . .

　　b.　*Wo zuotian　　yudao #(na ge)　nansheng*
　　　　I　yesterday meet　　that CLF boy
　　　　'I met the boy yesterday.'

　　c.　*Wo dai　gei #(na ge)　nansheng yi　ge　liwu*
　　　　I　bring give　that CLF boy　　　one CLF gift
　　　　'I'm bringing a gift for the boy.'

　　d.　*(na ge)　　nansheng kan-qi-lai you　er-shi　sui　zuoyou.*
　　　　that CLF boy　　　look　　have two-ten year or-so
　　　　'The boy looks twenty-years-old or so.'

　　e.　*Wo bu　renwei ?(na　ge)　nansheng hen　youqu.*
　　　　I　NEG think　that CLF boy　　very interesting
　　　　'I don't think that the boy is very interesting.'

---

[8]I will avoid the term *familiar* below. The term is attractive to familiarity theorists because of how it flexibly extends to accommodate examples like larger situation definites whose existence or uniqueness can be said to be taken for granted, hence licensing their familiarity in any discourse (Heim 1982). As such, the term seems ill-suited to a description which takes prior mention as a necessary condition on the use of this category.

[9]In most simple anaphoric environments I have checked, Mandarin speakers prefer *na* 'that' to *zhe* 'this.' Oshima and McCready (2016) show that in Japanese and English, proximal anaphoric demonstratives imply speaker-privileged familiarity with the referent, while distal demonstratives imply shared familiarity. The same basic contrast seems to be at play in Mandarin, explaining speaker preferences for *na*.

[10]Li (2013, 116-121) makes a similar observation about the distribution of bare nouns, claiming that unique definite interpretations are only available in object positions while anaphoric definite interpretations are only available in subject and topic positions. Yet bare nouns can also occur as unique definites in subject position ((11), (12b)).

Example (16b) shows that a bare anaphoric definite is judged infelicitous in object position, and (16c) shows that they are infelicitous as indirect objects.[11] Examples (16d) and (16e) illustrate that both bare nouns and demonstratives can occur in subject positions, both matrix and embedded, although most speakers report a preference for the demonstrative.[12]

Anaphoric definites in Mandarin must include demonstrative determiners even when the identity of the referent is unknown. In other words, demonstrative descriptions can refer *de dicto* provided an appropriate context:

(17)   MANDARIN: ANAPHORIC REFERENCE *de dicto*
    a.   *you ge nuren sha le Lisi*
        have CLF woman kill PRF Lisi
        'A woman killed Lisi.'
    b.   *jingcha huaiyi <u>na ge</u> <u>nuren</u> nashihou shou le shang*
        police   suspect that CLF woman at-that-moment suffer PRF injury
        'Police suspect that the woman suffered an injury.'
    c.   *jingcha huaiyi <u>nuren</u> nashihou shou le shang*
        police   suspect woman at-that-moment suffer PRF injury
        'Police suspect that a woman suffered an injury.'

In (17b), the demonstrative description refers back to the mysterious murderer in the first clause. (17c) illustrates that a bare noun in the same environment receives an indefinite interpretation, and cannot refer back to the murderer.

To conclude, Mandarin shows a general requirement for demonstratives with anaphoric definite noun phrases, with the exception of subject positions, which also allow a bare noun.

3.3   Donkey definites in Mandarin

Demonstratives are also required in Mandarin when noun phrases occur as donkey anaphora, anaphoric definites which receive quantificationally bound interpretations despite the absence of a c-commanding antecedent. While much attention historically has focused on donkey pronouns, recent work has focused on interpretation of definite descriptions (Elbourne 2005) and demonstratives (Abbott 2002) in donkey anaphoric environments:

(18)   a.   If a farmer has a donkey, he beats <u>the donkey</u>.
      b.   If a farmer has a donkey, he beats <u>that donkey</u>.

---

[11]Earlier work looking at this contrast (notably Jiang 2012) did not look at environments besides matrix subjects. This is why I believe the prohibition on anaphoric bare nouns in non-subject positions has not been previously noticed.

[12]Jiang (2012, ch. 4) shows that the facts are different for plural human nouns, due to the availability of the plural human suffix *-men* and the universal quantifier *dou*, both of which have been associated with definiteness, an issue I must set aside. I have found that non-human plurals, on the other hand, behave like singular nouns in allowing anaphoric definite interpretations of bare nouns in subject position, but not other positions, showing that this phenomenon is quite general.

This section shows that donkey definites, like other anaphoric definites in Mandarin, require demonstratives and prohibit bare nouns.

Mandarin has two types of donkey sentences: bare conditionals and *ruguo* or *dou*-conditionals (Cheng and Huang 1996). Bare conditionals only make use of indeterminate pronouns, so they are of little interest to us here. On the other hand, *ruguo* or *dou*-conditionals require a 'definite expression' in the consequent:

(19)   DOU-CONDITIONALS IN MANDARIN (Cheng and Huang 1996, ex. (22b,d))

       a.   *ni   jiao shei jin-lai, wo dou jian <u>ta</u>.*
            you ask who enter, I   all see him/her.
            'Whoever you ask to come in, I'll see him/her.'

       b.   *ni   jiao shei jin-lai, wo dou jian <u>na</u> <u>ge</u>   <u>ren</u>.*
            you ask who enter, I   all see that CLF person.
            'Whoever you ask to come in, I'll see that person.'

The relevant reading of (19) is one where the pronoun or demonstrative is bound, or where the choice of invitee covaries with the person who will be seen.

While Cheng and Huang (1996) observe that the class of 'definite expressions' which can serve as donkey anaphora in *dou* and *ruguo* conditionals include demonstrative descriptions and overt pronouns, they do not notice that bare nouns are impossible in this environment:

(20)   *#ni   jiao shei jin-lai, wo dou jian <u>ren</u>.*
         you ask who enter, I   all see person

This sentence has possible, but odd, interpretation in which the object of the main clause is interpreted as a low-scope indefinite. Thus, this sentence could only be translated as the bizarre *Whoever you ask to enter, I will see a person*.

The same restriction obtains if the donkey sentence is of the relative clause variety:

(21)   *mei   ge   [you yi zhi shuiniu de] nongfu dou hui da #(<u>na</u> <u>zhi</u>) <u>shuiniu</u>.*
        every CLF have one CLF buffalo REL farmer all will hit that CLF buffalo
        'Every farmer that has a buffalo hits that buffalo.'

Again, the bare noun in (21) can be interpreted generically, roughly equivalent to the bare plural object in English *Every farmer that has a donkey beats donkeys*.

So we see that the constraint on bare donkey definites is quite general in Mandarin. The observation that demonstratives (and overt pronouns) can occur as donkey anaphora in Mandarin while bare nouns cannot falls under the more general observation that demonstratives (and overt pronouns) can occur as anaphoric definites but bare nouns cannot. Together, narrative sequences and donkey sentences show that definite bare nouns are restricted to environments licensed by uniqueness.

4 Unique and anaphoric definites

This section presents an analysis of the contrast between bare nouns and demonstratives in Mandarin which builds on two ideas. First, I adopt an analysis of definite bare nouns in Mandarin via the type-shifting operator ι (Chierchia 1998; Yang 2001; Dayal 2004, 2011; Jiang 2012). Second, I adopt Schwarz (2009)'s account of weak versus strong definites in German, which are distinguished by an index just in the case of strong, anaphoric definites. I discuss the interplay between these two options in the following section.

4.1 Preliminaries

I assume that definite descriptions are individual denoting expressions of type e (Heim 1982, 1991; Elbourne 2013). I also adopt the semantics for common nouns and numeral classifiers of Trinh (2011), which builds on Chierchia (1998) and Krifka (1995). These proposals are based on the assumption that noun phrases are comprised of at least three distinct nominal projections: DP>ClfP>NP.

The semantic model contains a universe of discourse U which is made up of both individuals and pluralities (e.g. Link 1983; Schwartzschild 1996). Nouns are cumulative predicates consisting of both individuals and pluralities, closed under a sum operator $+$.[13] The universe of discourse must also include kind-level individuals and pluralities (cf. Dayal 2004). Classifiers (=Clf) are modeled as functions from cumulative predicates to atomic predicates (Chierchia 1998), where atomic predicates are essentially predicates which contain only individuals in their extension. In addition, classifiers serve to restrict the predicate denoted by the noun, which will range over both kinds and objects, to just one of these domains (Liao and Wang 2011; Nomoto 2013):[14]

(22) a. $x \in AT(P)$ iff $x \in P \wedge \forall y((y \in P \wedge y \leq x) \rightarrow (y = x))$ Atoms of P
       b. X is an atomic predicate iff $[\![X]\!]_s = AT([\![X]\!]_s)$
       c. $[\![N]\!] = \lambda x.\lambda s.P(x)(s)$
       d. $[\![Clf_{obj}]\!] = \lambda P.\lambda x.\lambda s.[P(x)(s) \wedge AT_{obj}(x)]$
       e. $[\![Clf_{kind}]\!] = \lambda P.\lambda x.\lambda s.[P(x)(s) \wedge AT_{kind}(x)]$
       f. $[\![Clf\ N]\!] = \lambda x.[P(x)(s) \wedge AT_{kind/obj}(x)]$

Finally, I will assume a situation semantics, which takes the existence of situations as semantic variables as a primitive (Barwise and Perry 1983; Kratzer 1989). In addition to serving the tradition roles of worlds in intensional contexts, situations serve as a domain restriction on determiners (Elbourne 2005, 2013; Schwarz 2009, 2012). Situations include

---

[13]That nouns in classifier languages are cumulative, i.e. number-neutral predicates, is defended at length in Rullman and You (2006) for Mandarin.

[14]Numerals will not be incorporated into this analysis. Krifka (1995) takes classifiers to be measure functions which require a numeral argument, presumably saturated by a silent 'one' when no numeral is pronounced. An alternative would be to analyze numerals as having their own measure function, but one which is only compatible with atomic predicates, effectively requiring the classifier.

partial or minimal situations, which can be made up of just an individual and a few *particulars*, contextually relevant properties of that individual.

4.2   Definite structures and meanings

Schwarz (2009, 2013) proposes that the difference between unique and anaphoric definites is that that anaphoric definites take an index as an argument while unique definites do not. However, both unique and anaphoric definites presuppose the existence of a unique individual (or a maximal plurality) to which they refer. The existence and uniqueness presuppositions hold within the context of a particular (minimal) situation (Heim 1990; Elbourne 2005, 2013), modeled as an argument of the determiner, a resource situation s_r, which functions as the domain restriction on the definite determiner. Anaphoric definite articles take one argument more than their unique counterparts, which is most contexts is satisfied by an index. My proposed denotations for the two types of definite articles, which I will abbreviate ι and ι^x, are provided below:

(23)   a.   UNIQUE DEFINITE ARTICLE
$$[\![\iota]\!] = \lambda s_r.\lambda P_{\langle e,\langle s,t\rangle\rangle}.: \exists!x[P(x)(s_r)].\iota x P(x)(s_r)$$
b.   ANAPHORIC DEFINITE ARTICLE: ι^x
$$[\![\iota^x]\!] = \lambda s_r.\lambda P_{\langle e,\langle s,t\rangle\rangle}.\lambda Q_{\langle e,t\rangle}.: \exists!x[P(x)(s_r) \wedge Q(x)].\iota x[P(x)(s_r)]$$

I have departed from Schwarz (2009, 2013) in the denotation of ι^x in that its indexical argument is a property, following the analysis of English demonstratives in Nowak (2014) and some other analyses of domain restrictions of presuppositional determiners (e.g. von Fintel 1994). Consequently, the index is only interpreted as part of the presuppositions of the anaphoric definite DP. I will call this argument the domain restriction below. The domain restriction can either be satisfied either by indices, which I take to be of type $\langle e, t\rangle$—following a suggestion of Elbourne 2005, ch. 3[15]— or other properties which provide contextual domain restrictions. Note that this property lacks a situation argument, unlike the nominal restriction of the determiner. Evidence and further discussion of these claims will be discussed in Section 4.4.

We turn now to the syntactic differences between unique and anaphoric definites in Mandarin. We have established that unique definites are realized as bare nouns, while anaphoric definites in most contexts require [Dem-Clf-N] phrases. Chierchia (1998) proposes that languages without overt definite articles like Mandarin make use of a semantic type-shifting operation to produce definite meanings. The definite type-shifter ι is just one of three such operators, but it is the only one that is immediately relevant.

To account for the restricted distribution of unmarked definite interpretations across languages, type-shifting operations are subject to a Blocking Principle:[16]

---

[15]One appealing aspect of this approach that person features, which are indisputably indices, are naturally modeled as predicates rather than individuals e.g. $\lambda x.[\text{SPEAKER}(x)]$.

[16]See Dayal (2004, 2011) and Jiang (2012) for additional arguments for this view.

(24)   *Blocking Principle:*
        Don't do covertly what you can do overtly!

The principle blocks definite type-shifting in languages with overt definite articles, forcing the projection of DP.

I will adopt this general approach, and further claim that Chierchia's definite type-shifter lacks an indexical argument so has the same denotation as a unique definite article ι in (23a). The fact that type-shifting cannot produce the meaning in ι^x is unsurprising: while a situation can be provided by the semantics, I assume that the domain restriction of ι^x must be present syntactically, a point I provide evidence for below. Because no domain restriction would be present if a bare noun was type-shifted with ι^x, such derivations would always crash. This leaves ι as the only type of definite interpretation which is available by type-shifting, which is possible in Mandarin precisely because it lacks a unique definite article.

On the other hand, I assume that the semantics for ι^x in (23b) is the regular semantics of Mandarin demonstratives. Both the nominal restriction and the domain restriction of demonstratives and other anaphoric definites must supplied syntactically because they are a part of its lexical meaning, although they might be deleted or silent in some cases. The domain restriction argument is represented as a DP adjunct. In the case of an index this argument is a null pronoun, which will be type-shifted to a property (again, see Section 4.4). For both ι and ι^x, the situation argument is supplied pragmatically.[17]

(25)   a.   UNIQUE DEFINITE   b.   ANAPHORIC DEFINITE
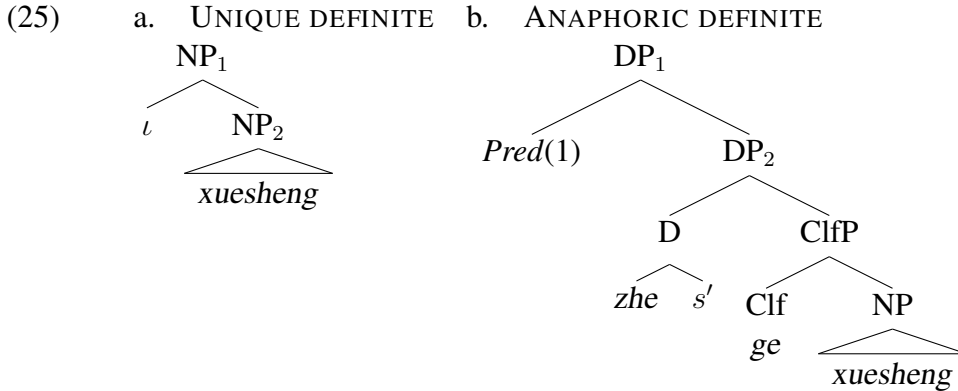
The interpretations of these two structures are provided below, building on the lexical entries in (22) and (23). Note that the index 1 is interpreted as an indexical property relativized to an assignment function: $\lambda x[x = g(1)]$.

(26)   Unique definite semantics ($[\![(25)\text{-a}]\!]$)
        a. $[\![\text{NP}_2]\!]^g = \lambda x.\lambda s.[\textbf{student}(x)(s)]$
        b. $[\![\text{NP}_1]\!]^g = \exists!x[\textbf{student}(x)(s')].\iota x[\textbf{student}(x)(s')]$

(27)   Anaphoric definite noun phrase ($[\![(25)\text{-b}]\!]$)

---

[17]See Schwarz (2012) for arguments that situation pronouns are arguments of determiners, specifically.

a. $[\![\text{NP}]\!]^g = \lambda x.\lambda s.[\textbf{student}(x)(s)]$

b. $[\![\text{ClfP}]\!]^g = \lambda x.\lambda s.[\textbf{student}(x)(s) \wedge AT_{\text{obj}}(x)]$

c. $[\![\text{DP}_2]\!]^g = \lambda Q_{\langle e,t \rangle}.\exists! x[\textbf{student}(x)(s') \wedge AT_{\text{obj}}(x) \wedge Q(x)].\iota x[\textbf{student}(x)(s') \wedge AT_{\text{obj}}(x)]$

d. $[\![\text{DP}_1]\!]^g = \exists! x[\textbf{student}(x)(s') \wedge AT_{\text{obj}}(x) \wedge x = g(1)].\iota x[\textbf{student}(x)(s') \wedge AT_{\text{obj}}(x)]$

The following two sections discuss how these interpretations account for the distribution of bare nouns and demonstratives definites in Mandarin, and provide additional support for the domain restriction argument of demonstratives in particular.

4.3 Unique definites and situations

The semantic contribution of the situation variable is an important component of the meaning of unique definites. In particular, there are contexts where a unique definite refers to different individuals as the choice of situation changes. Because ι is relativized to situations, we expect expressions involving ι to pick different individuals as choice of individual changes. In addition, the index in ι^x will block covarying readings in these same environments. This prediction is confirmed in the following examples (based on Elbourne 2005:21):

(28) SITUATIONALLY DEPENDENT REFERENCE IN MANDARIN

    a. *jin nian zongtong lai zi PFP*
       this year president come from PFP
       'This year [the president]$_i$ comes from the PFP.'

    b. *ming nian zongtong jiang shi DPP de dang yuan*
       next year president will be DPP REL party member
       'But next year [the president]$_{??i/j}$ will be from the DPP.'

    c. *ming nian zhe wei zongtong jiang shi DPP de dang yuan*
       next year this CLF president will be DPP REL party member
       'But next year [the president]$_i$ will be from the DPP.'

When the topic is quantificational, bound definite readings of bare nouns in Mandarin are possible (based on Schwarz 2009, ex. (231)):

(29) QUANTIFICATIONALLY BOUND SITUATIONS IN MANDARIN

    Obama mei dao yi ge chengshi ta dou gen (#zhe wei) shizhang ji anmi an
    obama every arrive one CLF city, he all with this CLF mayor meet
    'In every city that Obama visited, he met with the mayor (of that city).'

Examples (28b) and (29) illustrate that Mandarin bare nouns can receive covarying or sloppy interpretations, in which the president or mayor is different in different years.[18]

---

[18] 'Mayor' in (29) can receive a strict interpretation if there is a personal acquaintance of the speaker who is

In contrast, the demonstrative description in (28c) must receive a strict reading, one that is anaphoric to the president in the first sentence, and a demonstrative is infelicitous in the quantificationally bound example (29).

The semantic representation of covarying readings under situation binding for ι is illustrated in the following semantic paraphrases for (28b) and (29). Because the topical adverb supplies each sentence with a distinct situation, ι can return different individuals in each situation:

(30) a. This year=s_1, the unique president who is part of s_1 is a Republican. Next year=s_2, the unique president who is part of s_2 will be a Democrat.
      b. In every s, such that Obama visited a city in s, there is an s' that is part of s such that Obama visited the unique mayor who is part of s'

These meanings are approximate and gloss over several important details about the semantics of situations and the operators needed to derive such covarying readings; see Elbourne (2005, 2013) and Schwarz (2009) for details. Additionally, the sloppy readings of larger situation definites like 'president' and 'mayor' in the examples above are available because of general world knowledge about presidents and mayors: first, that these roles are unique at any particular time, and second, that these roles vary as a function of time.

Situation-based variation can also give rise to covarying readings of part-whole bridging definites because of world knowledge. Consider the Mandarin example below:

(31) *mei ge mai le fangzi de ren   dou xuyao xiuli (#na ge) wuding*
      every CLF buy PRF house de people all   need   fix   that CLF roof
      'Everyone that bought a house needed to fix the roof.'

Here, choice of roof varies with choice of house. The covarying reading is available because there is usually a unique roof that is part of any home-buying event.

With immediate situation definites, on the other hand, existence and uniqueness presuppositions are satisfied only relative to a topic situation which is part of the common ground:

(32) *Gou yao guo malu.*
      dog want cross road
      'The dog(s) want to cross the road.'   (Cheng and Sybesma 1999:510)

The bare noun *gou* 'dog' in subject position can receive three readings in (32), one generic, which we can set aside, and both a singular and plural definite reading. What is crucial is that the singular interpretation is only available in a situation where there is only one dog, and a plural interpretation must hold when there are multiple dogs.

---

a mayor and is called 'mayor.' A similar requirement holds for the strict interpretation of 'president' in (28b). I take this acquaintance condition to provide evidence for a directly referential use of these nouns akin to proper names. This directly referential use seems available to many human nouns in Mandarin and Cantonese, particularly titles and kinship terms.

4.4 Anaphoric definites as indexical expressions

Anaphoric definite environments such as narrative sequences (16) and donkey sentences (21), provide the classic motivation for dynamic theories of definiteness. In such theories, definite descriptions are interpreted as variables (Kamp 1981; Heim 1982; Groenendijk and Stokhof 1990, 1991), here one component of a more complex meaning. For Schwarz (2009) and the proposal adopted here, the indexical component of anpahoric definites is served by the domain restriction of ι^x when filled by an index, which can be interpreted dynamically.

Straightforward empirical support for the idea that the domain restriction of ι^x can be satisfied by a pronoun or something like it comes from the observation in Huang et al. (2009) that Mandarin allows pronouns (33) and proper names (34) can occur overtly before demonstrative descriptions, which serve to specify their reference (the free translations below are mine, under consultation):

(33)   PRONOUN + DEMONSTRATIVE IN MANDARIN (Huang et al. 2009, 298)

    a.   *wo xihuan [nimen zhe-xie guai haizi].*
        I   like   you.PL these   good children
        'I like you good kids.'
    b.   *wo dui [tamen naxie liulanghan] meiyou yinxiang.*
        I   to   they   those vagrant   not-have impression
        'I do not have impressions of them/those vagrants.'

(34)   PROPERN + DEMONSTRATIVE IN MANDARIN (Huang et al. 2009, 299)

    a.   *wo xihuan [Zhangsan, Lisi na   ji-ge   guai haizi].*
        I   like   Zhangsan, Lisi those several-CLF good children
        'I like those good kids like Zhangsan and Lisi.'
    b.   *wo dui [Zhangsan zhe-ge   xuesheng] mei-you shenme yinxiang.*
        I   to Zhangsan this-CLF student   not-have what   impression
        'I do not have much [of an] impression of this student Zhangsan.'

In these examples, the pronouns and names restrict or specify the intended reference of the noun phrase which they attach to. Under the proposal being adopted here, this is because they are functioning as the domain restriction of ι^x.

These expressions are not particularly exotic: but the proper name examples in (34) closely resemble close appositives in English:

(35)   a.   the poet Shakespeare
    b.   Shakespeare the poet

Or maybe, more equivalently to an example like (34b):

(36)   that guy John

Where postnominal *John* is the equivalent of the pre-determiner names in Mandarin. In-

terestingly, close appositives like those above are restrictive (e.g. Lekakou and Szendroi 2007), a point which is compatible with their analysis as domain restrictions for a determiner.[19]

The pronouns in (33) and the proper names in (34) must be interpretable as properties so that they can supply a domain restriction to ι^x, which is of type $\langle e, t \rangle$ in (23b). This can be achieved by applying the type shift *Pred* (Partee 1986, 1987) to domain restrictions of type e:

(37) a. *Pred*$(x)=$
    (i) $\lambda y[y = x]$ if $x \in D_e$
    (ii) Otherwise, undefined
    b. *Pred*$([\![tamen_4]\!]^g) = \lambda y[y = g(4)]$
    c. *Pred*$([\![Zhangsan]\!]) = \lambda y[y = \mathbf{z}]$ $\mathbf{z}$ = Zhangsan

This semantics is not quite right for plural examples like (34a), where the indexical arguments must simply be members of the plural referent of the DP, but the equivalence relation in (37) could be changed to a subset relation: *Pred*$(x) = \lambda y[x \subseteq y]$, where x must be contained in y. This would be trivially satisfied in singular examples.

The output of *Pred*, lacks a situation argument, which accounts for its suitability as domain restriction to ι^x. The fact that indexical agument Common nouns are prohibited in the pre-demonstrative position:

(38) **wo xihuan xuesheng na   liang-ge (ren)*
    I   like   student   that two-CLF person
    'I like those two students.' (intended)   (Huang et al. 2009, 301)

Common nouns are of type $\langle e, \langle s, t \rangle \rangle$, so they are semantically distinct from type-shifted indices.[20]

An analysis of example (34b) which makes use of the *Pred* typeshift is provided below:

(39) a.
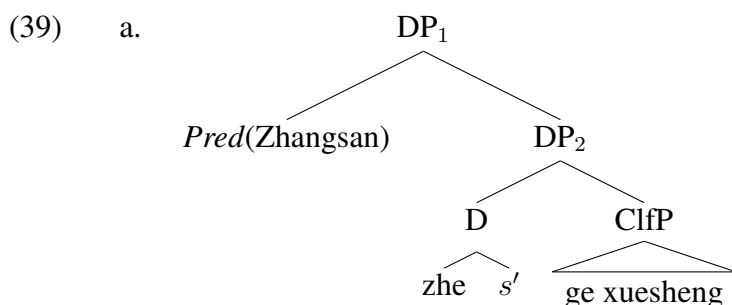
[19]Huang et al. (2009, 303-306) convincingly show that that the expressions in (33) and (34) do not consist of multiple appositive nominal expressions, but rather a single complex DP. This conclusion is compatible with the analysis advocated in this paper.

[20]Elbourne (2005, ch. 3) contemplates syntactic subcategorization as a way of preventing (type-shifted) indices from serving as the nominal restriction of determiners, but I take them to be semantically distinct.

b. $[\![\text{ClfP}]\!] = \lambda x.\lambda s[\textbf{student}(x)(s) \wedge AT(x)(s)]$

c. $[\![\text{DP}_2]\!]^g = \lambda Q.\exists!x[\textbf{student}(x)(s') \wedge AT_{\text{obj}}(x) \wedge Q(x)].\iota x[\textbf{student}(x)(s') \wedge AT_{\text{obj}}(x)]$

d. $[\![\text{DP}_1]\!]^g = \exists!x[\textbf{student}(x)(s') \wedge AT_{\text{obj}}(x) \wedge x = \textbf{z}].\iota x[\textbf{student}(x)(s') \wedge AT_{\text{obj}}(x)]$

In summary, the ability of names and pronouns to occur before demonstratives in Mandarin provides a relatively straightforward argument for the idea that ι^x takes an index as a syntactic argument, although it must be shifted to a predicative interpretation.

To this point, it has been more of an obstacle that the domain restriction of ι^x is *predicative*, but good evidence for this position comes from the availability of modifiers in the pre-demonstrative position in Mandarin, as in (40b) (Zhang 2015, ex. 1):

(40)   a.   *na   san ge   [mai le   ditan   de] ren*
               that three CLF buy PFV carpet REL person
               'those three people who bought a carpet'

        b.   *[mai le   ditan   de] na   san ge   ren*
               buy PFV carpet REL that three CLF person
               'those three people who bought a carpet'

Two restrictions hold for the pre-demonstrative modifier: it must be restrictive (Constant 2011), and it must be predicative (Zhang 2015). These restrictions follow from the semantics I am pursuing here. Additionally, the demonstrative in (40b) seems to lose its locative indexical (i.e. distal) semantics in the presence of a modifier. This is unsurprising if the usual null indexical domain restriction of the demonstrative, presumably accompanied by a distal presupposition, is blocked by the presence of an overt domain restriction in such examples.[21]

The same effect is found in English complex demonstratives (41), as noted by Nowak (2014), as well as demonstratives modified by locative pronouns, as in (42):

(41)   That guy who wrote *Waverly* was Sir Walter Scott.

Like its Mandarin counterpart, the complex demonstrative in (41) is odd in the context of ostension without focus on the demonstrative.[22] I conclude that in both English and

---

[21] See del Gobbo (2003), Lin (2003), and Boškovič and Hsieh (2012) for related discussions of high relative clauses and other modifiers as potential domain restrictions of the determiner.

[22] An anonymous reviewer points out the following example in Italian, with postnominal modifier *qui*, lacks a bound reading:

(42)   *Se Marco possiede un asino$_j$, quest' asino$_j$ (*qui)   mangia assia poco.*
           If Marco own   a donkey, this   donkey (over.here) eats   very little
           only meaning with qui: if M has a donkey, this (other) donkey over here eats little

This seems equivalent to the English examples below:

(43)   a.   That guy there was Sir Walter Scott.
        b.   Every man who had [a donkey]$_i$ beat [that donkey (*there)]$_i$/%[that there donkey]$_i$.

Mandarin, relative clauses may be fulfilling the semantic role of an index for ι^x, supplying the demonstrative with a domain restriction and blocking overt ostension to supply the index.

Having established that there is good evidence for an indexical argument in Mandarin, we turn now to anaphoric contexts, where this argument is saturated by a null index. We begin with the the cases of strict identity from the 'president' examples in (28). In examples like these, it is the presence of an index which enforces a *de re* discourse anaphoric interpretation. This effect can be attributed to the normal semantics of indices as variables, interpreted relative to an contextually provided assignment function g. A paraphrase of the semantic analysis of (28) with ι^x:

(44)   This year$=s_1$, the unique president in s_1 is a Republican. Next year$=s_2$, [the unique president in s_2 *identical to* $g(1)$] will be a Democrat.

The italicized identity condition in (44) above enforces a strict *de re* interpretation because the assignment function g is a constant parameter of interpretation in a particular context. The result is that the unique president in s_1 and s_2 must be one and the same person, whoever assignment function returns for the index 1.

Another semantic effect of the indexical argument of ι^x is to allow covarying readings in donkey sentences like (21). The fact that an index is required to derive such readings seems to support dynamic approaches to donkey anaphora, which rely on the presence of indices to derive these readings.[23]

There are two means by which dynamic theories derive covarying readings in these contexts (Chierchia 1995). For example, in Discourse Representation Theory these readings arise due to the semantic rule of unselective binding, where the free variables introduced by noun phrases (45a) are closed under universal quantifiers (45b) (Kamp 1981; Heim 1982; Kamp and Reyle 1993):

(45)   DRT INTERPRETATION OF *Every man who owns a donkey beats it.* (simplified)
    a. $[[\mathbf{man}(x) \wedge \mathbf{donkey}(y) \wedge \mathbf{owns}(x,y)] \rightarrow \text{beat}(x,y)]$
    b. $\forall x \forall y [[\mathbf{man}(x) \wedge \mathbf{donkey}(y) \wedge \mathbf{owns}(x,y)] \rightarrow \mathbf{beat}(x,y)]$

Dynamic Predicate Logic uses somewhat different mechanisms, defining special dynamic connectives that result in cross-clausal binding (Groenendijk and Stokhof 1991). The effect of these connectives is that the scope of existential quantifiers can extend out of the clauses in which they are contained, allowing the indefinite in relative clauses (46a) to scope into the consequent as in (46b):

(46)   DPL INTERPRETATION OF *Every man who owns a donkey beats it.* (simplified)
    a. $\forall x [\mathbf{man}(x) \wedge \exists y [\mathbf{donkey}(y) \wedge \mathbf{owns}(x,y)]] \mathbf{beat}(x,y)$

In both cases, the locative elements must be saturating the extra index, blocking a bound reading.

[23] See the closely related argument by Schlenker (2011) for dynamic binding based on the requirement for indexical expressions in donkey sentences in two sign languages.

b. $\forall x[[\textbf{man}(x) \wedge \exists y[\textbf{donkey}(y) \wedge \textbf{owns}(x,y)]] \rightarrow \exists y[\textbf{donkey}(y) \wedge \textbf{owns}(x,y) \wedge \textbf{beat}(x,y)]]$ (Chierchia 1995, 124)

Despite these differences, both theories assume that donkey anaphora arise due to binding of an index. The expressions below plug the at-issue contribution of DPs headed by ι^x into the bound position in each of the two analyses above:

(47) Covarying reading via ι^x in Discourse Representation Theory
$\forall x \forall y[[\textbf{man}(x) \wedge \textbf{donkey}(y) \wedge \textbf{owns}(x,y)] \rightarrow \textbf{beat}(x, \underline{\iota z[\textbf{donkey}(z) \wedge AT_{\text{obj}}(z) \wedge z = y]})]$

(48) Covarying reading via ι^x in Dynamic Predicate Logic
$\forall x[[\textbf{man}(x) \wedge \exists y[\textbf{donkey}(y) \wedge \textbf{owns}(x,y)]] \rightarrow \exists y[\textbf{donkey}(y) \wedge \textbf{owns}(x,y) \wedge \textbf{beat}(x, \underline{\iota z[\textbf{donkey}(z) \wedge AT_{\text{obj}}(z) \wedge z = y]})]]$

We can see in these examples that the indexical argument of ι^x, occupied by the variable y, is bound. If the DP in these environments was headed by ι, no index would be available for binding and a covarying interpretation might not obtain.

5 Competition between definite expressions

The earlier sections have provided evidence for the semantic distinction between unique and familiar definites in Mandarin and proposed a syntax and semantics both types of definites. This section offers solutions to a few remaining puzzles, concerning the full distribution of definite expressions in Mandarin. First, it is somewhat unclear why ι^x is unavailable without prior mention. Second, it is similarly unclear why ι is impossible in anaphoric environments. Finally, it is unclear why optionality between the two definite expressions emerges in subject position. In this section I offer explanations for these restrictions.

5.1 On the availability of indices

In Mandarin, demonstratives are not allowed in unique definite environments. We saw that this was true even in cases of covariation like (29) and (31). Together, these facts provide a clear indication that ι^x is unavailable in unique definite environments. Why might this be?

Recall that unique definite environments are distinguished from anaphoric definite environments in that they do not involve prior mention in the discourse. A likely explanation for the unavailability of ι^x in unique definite environments, then, is that the index which distinguishes ι^x from ι is only licensed by explicit prior mention in discourse. In the one exception to this generalization, part-whole bridging, it is prior mention of an argument of the noun which licenses ι^x.

The prior mention condition on indices is evident with pronouns as well as the 'formal link' requirement, illustrated below (Heim 1982, 1991):

(49) a. Every man who has a wife is sitting next to her.
b. ?*Every married man is sitting next to her.

(50) a. Someone who has a guitar should bring it.
      b. ?*Some guitarist should bring it.

Elbourne (2001, 2005) shows that the formal link requirement receives a natural explanation in an ellipsis-based analysis of pronouns. Because ellipsis generally requires previous mention of the elided material to be licensed (Hankamer and Sag 1976; Merchant 2001), the formal link requirement falls out naturally if pronouns are D heads with a deleted NP complement. This account is extended to both personal and demonstrative pronouns in German by Patel-Grosz and Grosz (2017), where the formal link requirement also plays an important role.

If this proposal is on the right track, the prior mention requirement on indices might also be related to ellipsis licensing. Suppose, for example, that domain restrictions themselves must be elided when they are covert, whether they are pronouns or phrasal modifiers such as relative clauses. Such a proposal has an antecedent in Heim (1990)'s E-type analysis of donkey anaphora, which involve a hidden relative clauses copied onto noun phrases at LF.[24] The bound variable contained in this relative clause enables a covarying interpretation. In other words, a normal donkey sentence like (51a) would have a representation like (51b) at LF:

(51) a. Every farmer who owns a donkey beats <u>it</u>.
      b. [Every farmer who owns a donkey]$_j$ beats <u>[[it donkey]$_i$ that he$_j$ owns t_i]</u>.

This analysis is reminiscent of the recent proposal by Collins (2015) that domain restrictions for determiners are always explicit in the sense of Neale (1990), and are optionally deleted. Suppose that pronominal indices, which we saw can be overt in Mandarin (33), are also subject to ellipsis licensing. Then the prior mention requirement on ι^x can be reduced to ellipsis licensing as well.

5.2   On the unavailability of bare nouns in anaphoric environments

While anaphoric definites are never possible in unique definite environments, we saw in Section 3.2 that bare nouns, which are type-shifted via ι, are possible as anaphoric definites in subject position but not other syntactic positions. In the context of this analysis, two additional points need clarification. First, why are bare nouns infelicitous in anaphoric definite environments? Second, what exempts subjects from this restriction? This section addresses the first question. The suggestion is that bare nouns run afoul of a constraint called *Index!* which prefers to use indices when they are available, a specialized form of *Maximize Presupposition!* (Heim 1991).

The unavailability of bare nouns anaphoric environments in Mandarin mirrors similar facts in German. Specifically, Schwarz (2009, ch. 6) notes that weak article definites are unavailable without prior mention. Schwarz suggests that the problem with using ι in these

---

[24]Recall that LF-copying was the standard analysis of ellipsis until the advent of the Minimalist Program. Hence, this is essentially an ellipsis-based analysis of donkey anaphora.

environments that its uniqueness presupposition is not satisfied. In other words, the mere mention of some entity does not suffice to establish the uniqueness of such an entity, even if no other entity of the same sort has been introduced. Schwarz also suggests that the uniqueness presupposition of ι also fails in donkey sentences. Take, for example, a context where every man beats his respective donkey. In such a context, there would be multiple donkeys in the domain discourse, leading to a presupposition failure. In such contexts, the thinking is, an index is required in order to restrict the definite to a single donkey covarying with the choice of donkey-beaters.

If correct, this conclusion would suggest that a situation variable binding accounts of donkey anaphora (e.g. Elbourne 2005, 2013) are systematically unavailable to UG, a conclusion reached by Jenks (2015) based on similar facts in Thai. But this conclusion has a fatal flaw: definite bare nouns and German weak definites *can* receive covarying interpretations as long as they are not mentioned earlier, as we saw in (31), repeated below:

(52) *mei ge mai le fangzi de ren dou xuyao xiuli (#<u>na</u> ge) <u>wuding</u>*
      every CLF buy PRF house de people all need fix that CLF roof
      'Everyone that bought a house needed to fix the roof.'

If the problem with using ι in donkey sentences was the failure of its uniqueness presupposition, we would expect a bare noun to be unavailable in (52), given that there is no unique house in the domain of discourse. Yet these sentences require a bare noun, and their German counterparts require a weak article. So Schwarz (2009)'s explanation for the infelicity of ι in anaphoric environments based on the failure of uniqueness cannot be right, nor can Jenks (2015)'s conclusion about the unavailability of situation variable binding, which must be available to produce covarying readings in the example above (Section 4.3). In summary, we need another explanation for the infelicity of ι in anaphoric definite contexts besides the failure of uniqueness.

Upon further reflection, the availability of covarying readings for ι in just those environments where ι^x is not licensed provides a clue to the explanation for why ι is impossible in anaphoric environments. In particular, it suggests that there is a default preference in Mandarin and German for explicitly representing indices whenever possible:

(53) *Index!*
      Represent and bind all possible indices.

Because ι^x includes an index which is absent in ι, ι^x will be preferred whenever it is available. Crucially, because the index is part of the presupposition of ι^x, *Index!* is a specific instance of *Maximize Presupposition!* (Heim 1990), a principle which can in turn be reduced to the effect of a Gricean quantity implicature (Schlenker 2012). In this light we can conclude that the prediction of previous work is that the competition between definites be subject to *Index!*. Returning to the distribution of ι and ι^x, *Index!* has the effect of reducing ι to a kind of elsewhere determiner, only possible when no index is available due to the absence of prior mention. The prediction of a principle like *Index!* comports with

the facts, as in most contexts where indices are available in both Mandarin and German the demonstrative or strong determiner must be used.

5.3   Subjects, topics, and definiteness

Of course the one exception to the predictions of *Index!* is subject position, where we saw in Section 3.2 that bare nouns and demonstratives are in apparent free variation.

The first point I would like to make about this observation is that it is impossible to determine whether a similar effect obtains in German because of the details of the weak/strong article distinction there. In particular, the weak/strong article distinction is only apparent with objects of certain prepositions, because it conditions contraction. Because objects of prepositions by definition are not subjects, it is impossible to know whether articles in subject position are weak or strong. Thus, it could be that the generalization about the exceptionality of subjects in Mandarin is in fact quite general among languages that distinguish unique and anaphoric definites; we must understand the facts in more languages before we know.

In either case, the explanation I would like to pursue is based just on Mandarin. The main idea is that anaphoric bare nouns in subject positions are exceptional because they are continuing topics. The fact that Mandarin subjects are often topics has been well-established. For example, Cheng and Sybesma (1999) observe that subjects are almost always definite in both languages, an observation which clearly relates to their status as topics (Li and Thompson 1981).[25]

The basic idea is that the use of a weak definite such as a bare noun in subject position serves to mark a noun phrase as a topic. This pragmatic function of topic marking overrides and neutralizes the effect of *Index!* in such environments. The intuition behind this analysis is clear: topics do not need to be indexed because they are salient members of the Question Under Discussion (Roberts 1996; Büring 2003), modeled as a topic situation, following Schwarz (2009). In the context of the topic situation, the uniqueness presupposition of ι will always be sufficient to identify the intended referent. While non-subjects noun phrases might be part of the common ground, they are not topics. The reason for this is simple: non-subject topics in Mandarin are typically realized as (often null) pronouns (Huang 1984).

Initial evidence that bare nouns are pragmatically marked comes from the judgments of native speakers. While demonstratives and bare nouns were both judged to be felicitous in subject position, many speakers preferred demonstratives. This leads to the expectation that the bare noun might be pragmatically marked, as the topic reading would require additional pragmatic accommodation.

More specific evidence for the topicality of a bare nominal subject is provided by the question-answer pair in (54), where we have controlled to ensure a non-topical subject.[26] Adopting the general framework of Roberts (1996), we can say that Question Under Dis-

---

[25]See Yang (2001) and Jiang (2012) for a more refined discussion of subject definiteness and topicality in Mandarin.

[26]Thanks to an anonymous reviewer for suggesting this particular context.

cussion (QUD) introduced by (54A) is 'What happened to the mouse?' The mouse is the topic. While the cat named *xiao-hei* 'little black' has been named and is known, it is not the topic and is not included in the minimal topic situation. The answers below, one active (54B) and the other passive (54B′), show that anti-topical cat cannot be a bare noun even in subject position:

(54) A: *Zuihou na zhi bei   xiao-hei   zua-dao de lao-shu zenme   le?*
          finally that CLF PASS little-black catch   REL mouse what.happened PFV
          'What happened to the mouse that was caught by 'Little Black' (a black cat)?'
      B: #*(Na zhi) mao sha   le   (ta).*
          that CLF cat   killed PFV it
      B′: *Ta bei   #(na zhi) mao sha le*
          it PASS that CLF cat   kill PFV

This example clearly suggests that in the earlier examples where bare nouns could occur in subject position, such as in (16d), it was serving as a topic.

The kind of topic that can be marked with a bare noun must be a continuing topic. The topic cannot be a new topic, which might be marked with left-dislocation, extrapolating from work on English (Prince 1998). As such, we predict, correctly, that a left-dislocated topic would still prefer a demonstrative:

(55) a. *jiaoshi   li   zuo-zhe yi ge nansheng he yi ge nusheng,*
          classroom inside sit   Prog one CLF   boy one CLF girl,
          'There is a boy and a girl sitting in the classroom . . .
      b. #*(na ge) nansheng, wo hen   bu xihuan.*
          that CLF boy,   I   really not like
          'The boy, I really don't like.'

Now take a similar but subtly different discourse structure, that of a contrastive topic (Büring 2003), where a list of salient alternatives are being described relevant to some QUD. In particular, if we introduce the topic in narrative sequence above with the contrastive topic marker *ne* (Constant 2014), a bare noun is possible:

(56) *nansheng **ne**, wo hen   bu xihuan.*
      boy   CT, I   really not like
      'The boy, I really don't like.'

One implicature arising from the contrastive topic marker is that the speaker does like the girl in the context. This is because contrastive topics are continuing topics, but in such cases the QUD is about an alternative set rather than a single individual, e.g. 'How do you feel about boy x and girl y?.' This corresponds to the observation by Jiang (2012) that anaphoric bare nouns are licensed by contrast, which we can see as one member of a topical set of alternatives relevant to a particular QUD.

Much more ultimately needs to be said about the relationship between the shape of

discourse and the choice of definite, pronoun, or null anaphor in Mandarin. The main point here is that anaphoric bare nouns in subject position seem to mark continuing topics, a pragmatically marked use, accounting for speaker preference for the demonstrative in this position in elicitation contexts. The pragmatic function of topic-marking must take precedence over the effect of *Index!*.

6 Cantonese and the typology of definiteness marking

Not all languages without definite articles have the same distribution of definite expressions as Mandarin. This section describes the realization of definiteness in Cantonese, which allows [Clf-N] phrases in definite environments (Cheng and Sybesma 1999, 2005; Sio 2006). We will see that these phrases occur in both unique and anaphoric definite environments, and the logic of the previous analysis leads to the conclusion that Cantonese must have an ambiguous definite article, a conclusion which also applies to English *the*.

6.1 Definiteness in Cantonese

The examples below illustrate that [Clf-N] phrases in Cantonese occur in both unique and familiar definite environments:

(57)   UNIQUE DEFINITE CLF-N SEQUENCE IN CANTONESE
   a.   _Lou^{5}baan^2_ haa^{6}zau^3 lei^5 gim^{2}caa^4 gung^{1}zok^3.
         boss   afternoon come inspect   work
         'The boss is coming for an inspection this afternoon.'
   b.   _Zung^{2}tung^2_ gam^{1}maan^5 baat^3 dim^2 soeng^5 din^{6}si^6.
         president   tonight   eight o'clock get.on television
         'The president is going to be on TV at 8pm.'

(58)   NARRATIVE SEQUENCE WITH [CLF-N] PHRASE IN CANTONESE
   a.   _lei^{5}sei^3 fong^{2}man^6 zo^2 jat^1 go^3 zok^{3}gaa^1 tung^{4}maai^4 jat^1 go^3_
         Lei-sei interview   PERF one CLF writer   and   one CLF
         zing^{3}zi^{6}gaa^1.
         politician
         'Lei-sei interviewed a writer and a politician.'
   b.   _Keoi^5 m^4 jing^{6}wai^4 (#go^2) go^3 zing^{3}zi^{6}gaa^1 hou^2 jau^{5}ceoi^3._
         3SG   NOT think   that CLF politician   very interesting
         'He didn't think that the politician was very interesting.'

(59)   CANTONESE DONKEY SENTENCE WITH [CLF+N] PHRASE
        mui go jau jat zek maa ge lungfu daa zek maa.
        every CLF have one CLF horse REL farmer hit CLF horse
        'Every farmer that has [a horse]$_i$ hits [that horse]$_i$.'

27

While this is a small data set, additional discussion with Cantonese speakers has revealed that [Clf-N] phrases are systematically available in both unique and familiar definite environments. This means that [Clf-N]-phrases have the same distribution as definite articles in languages like English.

As a result, demonstratives and bare nouns cannot occur in these environments. These expressions have a more restricted distribution in Cantonese. In the case of demonstratives, they restricted to specific anaphoric environments, such as referring back to an individual who was mentioned several sentences ago, as in the sentences in (60c):[27]

(60) NARRATIVE SEQUENCE IN CANTONESE WITH MULTIPLE ANTECEDENTS

    a.   *kam^{4}jat^6 ngo^5 jing^{4}sik^1 zo^2 jat^1 go^3 hou^2 jau^5 mui^{6}lik^3 ge^3 bak^{1}ging^1*
        yesterday 1SG meet   PERF one CLF very have charm   POSS Beijing
        hok^{6}saang^1
        student]$_i$
        'Yesterday I met a charming student from Beijing.'

    b.   *jin^{4}hau^6 ngo^5 jau^6 jing^{4}sik^1 zo^2 jat^1 go^3 hou^2 jau^5 meng^2 ge^3*
        afterwards 1SG also meet   PERF [one CLF very have name POSS
        jing^{1}gwok^3 gaau^{3}sau^6
        England   professor]$_j$
        'Afterwards, I also met a famous professor from England.'

    c.   (i)   *go^2 go^3 hok^{6}saang^1 bei^2 go^3 gaau^{3}sau^6 cung^{1}ming^4*
             that CLF student   CMPR CLF professor smart
             'The student was smarter than the professor.'

        (ii)   *go^3 gaau^{3}sau^6 bei^2 go^2 go^3 hok^{6}saang^1 ceon^2*
             CLF professor CMPR that CLF student   stupid
             'The professor was dumber than the student.'

This specialized use of the demonstrative resembles the types of uses described for English in Wolter (2006) in that it is much more restricted that the Mandarin equivalents.

What about bare nouns? Simpson et al. (2011) observe that languages like Cantonese, which typically require classifiers in definite noun phrases, do sometimes allow definite bare nouns. Consider the following sentences:

(61)   a.   *Lou^{5}baan^2 haa^{6}zau^3 lei^5 gim^{2}caa^4 gung^{1}zok^3.*
        boss   afternoon come inspect   work
        'The boss is coming for an inspection this afternoon.'

     b.   *Zung^{2}tung^2 gam^{1}maan^5 baat^3 dim^2 soeng^5 din^{6}si^6.*
        president   tonight   eight o'clock get.on television
        'The president is going to be on TV at 8pm.'

The subjects in these sentences are putative instances of larger situations definites, which patterned with uniqueness definites (Section 3.1). As such, these bare nouns give the im-

---

[27]I am indebted to Herman Leung for this insight.

pression that Cantonese can use bare nouns in some definite environments.

However, there are some important restrictions on definite bare nouns in Cantonese. The first is that they are all arguably terms of address or reference, including kinship terms such as *ma¹ma¹* 'mother', *ba¹ba¹* 'father', and *jie³jie³* 'older sister', all nouns that would be used as titles, or terms of address in particular situations.[28] Additionally, when they are bare nouns, speakers judge that definites such as *lou⁵baan²* 'boss' and *zung²tung²* 'president' refer to an individual that the speaker and hearer both know personally.

I would like to suggest that these bare nouns in Cantonese have the semantics of proper names. Evidence for this conclusion comes from the fact that while a [Clf-N] phrase can covary in situationally covarying contexts (62), bare nouns do not (63):

(62)   *hai²*   *ou³baa¹maa²*   *heoi³*   *gwo³*   *ge³*   *mui⁵*   *jat¹*   *go³*   *sing⁴si⁵,*   *keoi⁵*   *dou¹*   *tung⁴*
        be.at   Obama   go   PFV   POSS   every   one   CLF   city,   s/he   all   with
        <u>*go³*</u>   <u>*si⁵zoeng²*</u>   *gin³*   *gwo³*   *min⁶.*
        CLF   mayor   see   PFV   face.
        'In every city that Obama visited, he met with the mayor (of that city).'

(63)   *hai²*   *ou³baa¹maa²*   *heoi³*   *gwo³*   *ge³*   *mui⁵*   *jat¹*   *go³*   *sing⁴si⁵,*   *keoi⁵*   *dou¹*   *tung⁴*
        be.at   Obama   go   PFV   POSS   every   one   CLF   city,   s/he   all   with
        <u>*si⁵zoeng²*</u>   *gin³*   *gwo³*   *min⁶.*
        mayor   see   PFV   face.
        'In every city that Obama visited, he met with the mayor.' (of some random city, 'the mayor' we are both acquainted with)

The lack of covarying readings with a definite bare noun in Cantonese is the opposite pattern from the Mandarin one, where a covarying reading was only allowed with bare nouns. As such, definite bare nouns in the two languages must be semantically distinct. Interesting, other classifier languages with Clf-N patterns have recently been shown to allow high-animacy exceptions as well (Simpson 2017; Simpson and Biswas 2016).[29] Such human bare nouns strike me as likely to also be a kind of 'common proper noun,' and are likely directly referential which cannot covary.

To summarize, then, [Clf-N] phrases in Cantonese occur in both unique and anaphoric definite environments, just like in a language with a general purpose definite article like English. The resulting typology of Cantonese, Mandarin, English, and German is summarized in Table 1.

Why does Cantonese express definites with Clf-N? The simplest answer is that Can-

---

[28]Of course the same phenomenon occurs with English kinship terms, but it is more restricted, occurring only with *mom*, *dad*, *grandpa*, and *grandma* (and their dialectal variants), but not *my brother*, *my cousin*, *my son*, etc.

[29]Interestingly, Simpson (2017); Simpson and Biswas (2016) observe that inanimate nouns occur as bare nouns in some environments where a definite would occur in definite article languages. Yet many of these environments resemble weak definites environments, another place bare nouns seem to occur in Cantonese. As weak definites are principally notable because they are a kind of 'fake' definite marking in a semantically indefinite environment, it is not clear to me that bare nouns in these environments are definite at all.

Table 1: DEFINITENESS MARKING IN CHINESE AND GERMANIC

	Mandarin	German	Cantonese	English
Unique definites (ι)	N	weak article	Clf-N	*the*
Anaphoric definites (ι^x)	Dem-Clf-N	strong article	Clf-N	*the*

tonese has a lexical definite article which licenses the classifier, like demonstratives do in all Chinese dialects and many other classifier languages such as Thai and Burmese. This null article might trigger Clf-to-D movement (cf. Simpson 2005; Wu and Bodomo 2009). Because a lexical article is available, type-shifting is unavailable, due to the Blocking Principle in (24).

In addition, we can conclude that definiteness marking in English and Cantonese must be ambiguous between ι and ι^x as long as *Index!* is universally active as a component of basic pragmatic reasoning. This is because the same definite form occurs in both unique and anaphoric contexts in English and Cantonese despite the availability of anaphoric demonstratives. If definite articles in these languages only exponed ι, *Index!* would require that a demonstrative take over in anaphoric contexts. We must conclude that these definite expressions can include indices in both languages. In other words, it is a consequence of the analysis developed here that English *the* must be ambiguous.

6.2 The typology of definiteness marking

Finally, we turn to a more extended typology of definiteness which supports both the ambiguity of English and Cantonese definite expressions as well as the idea that Mandarin definite bare nouns arise due to last resort type-shifting.

A cross-linguistic surveys of the cases of languages that distinguish two types of definiteness, most of which were discussed in Section 2, reveal that there are in principle four types of definiteness marking languages, only three of which actually occur. First, there are *bipartite* languages which have two separate articles for anaphoric and unique definites, such as Germanic languages and Lakhota (Schwarz 2013). Second, there are *marked anaphoric* languages which have a definite article which is restricted to anaphoric definite environments, including Fante Akan (Arkoh and Matthewson 2013) and some Wu Chinese dialects that restrict [Clf-N] phrases to anaphoric environments (Li and Bisang 2012; Simpson 2017). And third, of course there are *generally marked* definite languages like Cantonese and English, where a single definite form is used in both contexts.

However, there is a typological gap corresponding to a logically possible *marked unique* type of language, where a definite marker would occur in unique environments but not anaphoric ones. The complete typological picture is summarized in Table 2.[30]

---

[30]It is unclear if there are languages where definite bare nouns are possible in both unique and anaphoric

Table 2: TYPOLOGY OF DEFINITENESS MARKING

	Bipartite	*Marked anaphoric*	*Generally marked*	*Marked unique*
Unique (ι)	Def_{weak}	$\emptyset$	Def	Def_{weak}
Anaphoric (ι^x)	$\text{Def}_{\text{strong}}$	$\text{Def}_{\text{strong}}$	Def	$\emptyset$
Languages	German, Lakhota	Mandarin, Akan, Wu	Cantonese, English	(*unattested*)

There is a historical explanation for the absence of marked unique languages. Greenberg (1978) observes that definite articles typically are grammaticalized from demonstratives, and first show up in anaphoric contexts. Mandarin is at the starting point in this historical chain. However, it is impossible for a language to grammaticalize a unique definite without first having an anaphoric definite article. In the context of the typology in Table 2, the proposal is essentially that generally marked languages like English are semantically bipartite.

This typology is a powerful argument for the analysis of Mandarin proposed in this paper. Suppose bare nouns in Mandarin were due to the presence of a null article interpreted as ι, which would suggest that either of ι or ι^x can be null in different languages. We then clearly predict that null ι^x should be found in some languages with an overt ι, as this would essentially be a random phonological property of these heads. As a result, we would have no reason to expect the absence of marked unique languages. If the type-shifted ι and demonstratives are both universally available, Mandarin represents the most minimal type of definiteness marking language, making use only of generally available resources.

And to return to the opening theme, what does this tell us about the structures which realize definites in Mandarin and similar languages? It suggests that bare nouns lack DP structure in Mandarin, but that the presence of demonstratives involve full DP structure. As such, the debate about whether a language is a 'DP language' or an 'NP language' is not on quite the right track. Instead, we must ask under which semantic contexts languages are required to project DP.

7 Conclusion

In this paper I have shown that Mandarin distinguishes unique from anaphoric definites. While bare nouns occur in unique unique definite environments, demonstratives were re-

---

environments. The description of Hindi in Dayal (2004) suggests that this might be true, but only a single relevant example is given. Such languages would have to be analyzed as generally marked languages with a null definite determiner in my system.

quired in anaphoric environments.

No novel theoretical innovations were needed to account for these facts. The unique definite interpretation of bare nouns was attributed to the type-shifting operator ι, which has been proposed for bare nouns in Mandarin in previous work (Chierchia 1998; Yang 2001; Dayal 2004; Jiang 2012). As for demonstratives, their anaphoric use was seen as a special instance of their independently necessary indexical semantics, and several surprising facts about the distribution of pre-demonstrative modifiers in Mandarin were shown to follow from this proposal.

I also proposed relatively detailed accounts of how general grammatical principles mediate competition between the two forms. I argued first that the prior mention requirement on ι^x follows from ellipsis licensing, second that a principle *Index!* prefers ι^x whenever these conditions are met, and third that the exceptional uses of bare nouns as anaphoric definites in subject position is accounted for if we take such occurrences to be a form of topic marking.

The picture that emerges is one where Mandarin is a kind of default language in the typology of definiteness: what you get without any lexical article. This is different from a language like Cantonese, which despite the absence of an overt article does indicate its presence via a [Clf-N] phrase. We might expect that as a default, languages like Mandarin might be common — this is probably true. For example, Jenks (2015) shows that basically identical facts hold in Thai, Oshima and McCready (2016) describe the distribution of anaphoric demonstratives in Japanese, which also has definite bare nouns, and work on Korean (Lee 1995; Kang 2015; Kim and Yoon 2016) suggests the same distribution there as well. If these facts turn out to be fully general, we may be well on our way to a universally applicable theory of definiteness.

References

Abbott, Barbara. 2002. Donkey demonstratives. *Natural Language Semantics* 10:285–298.

Abbott, Barbara. 2010. *Reference*. Oxford: Oxford University Press.

Abney, Steven. 1987. The noun phrase in its sentential aspect. Doctoral Dissertation, MIT.

Arkoh, Ruby, and Lisa Matthewson. 2013. A familiar definite article in Akan. *Lingua* 123:1–30.

Barker, Chris. 2005. Possessive weak definites. In *Possessives and Beyond: Semantics and Synax*, ed. Ji yung Kim, Yury A. Lander, and Barbara H. Partee. Amherst, MA: GLSA Publications.

Barlew, Jefferson. 2014. Salience, uniqueness, and the definite determiner -tè in Bulu. In *Semantics and Linguistic Theory (SALT) 24*, ed. Todd Snider, Sarah D'Antonio, and Mia Weigand, 619–639.

Barwise, Jon, and John Perry. 1983. *Situations and attitudes*. Cambridge: MIT-Bradford.

Bisang, Walter. 1999. Classifiers in East and Southeast Asian languages: Counting and beyond. In *Numeral types and changes worldwide*, ed. Jadranka Gvozdanovic, 113–185. Berlin: Mouton de Gruyter.

Boškovič, Zeljko. 2008. What will you have, DP or NP? In *Proceedings of NELS 37*, ed. Martin Walkow and Emily Elfner.

Boškovič, Zeljko. 2012. On NPs and clauses. In *Discourse and grammar: From sentence types to lexical categories*, ed. Günter Grewendorf and Thomas Ede Zimmermann, 179–242. Berlin: Walter de Gruyter.

Boškovič, Zeljko, and I-Ta Chris Hsieh. 2012. On word order, binding relations, and plurality within Chinese NPs. In *Proceedings of the 13th International Symposium on Chinese Language and Linguistics*, ed. Yung-O Biq and Lindsey Chen.

Büring, Daniel. 2003. On d-trees, beans, and b-accents. *Linguistics and Philosophy* 26:511–546.

Carlson, Gregory, Rachel Sussman, Natalie Klein, and Michael Tananhaus. 2006. Weak definite noun phrases. In *Northe East Linguistic Society (NELS) 36*, ed. Chris Davis, Amy Rose Deal, and Youri Zabbal, 179–196. GLSA.

Chen, Ping. 2004. Identifiability and definiteness in Chinese. *Linguistics* 42:1129–1184.

Cheng, Hsu-Te. 2013. Argument ellipsis, classifier phrases, and the DP parameter. Doctoral Dissertation, University of Connecticut.

Cheng, Lisa L.-S., and C.-T. James Huang. 1996. Two types of donkey sentences. *Natural Language Semantics* 4:121–163.

Cheng, Lisa L.-S., and Rint Sybesma. 1999. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30:509–542.

Cheng, Lisa L.-S., and Rint Sybesma. 2005. Classifiers in four varieties of Chinese. In *The Oxford Handbook of Comparative Syntax*, ed. Guglielmo Cinque and Richard Kayne, 259–292. Oxford University Press.

Cheng, Lisa L.-S., and Rint Sybesma. 2012. Classifiers and DP. *Linguistic Inquiry* 43:634–650.

Chierchia, Gennaro. 1995. *Dynamics of Meaning*. Chicago: University of Chicago Press.

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

Clark, Herbert H. 1975. Bridging. In *Theoretical issues in natural language processing*, ed. P. N. Johnson-Laird and P. C. Wason. Association for Computing Machinery.

Collins, Chris. 2015. Quantifier domain restriction as ellipsis. NYU Ms., available at `ling.auf.net/lingbuzz/002616/`.

Constant, Noah. 2011. Re-diagnosing appositivity: Evidence for prenominal appositives from Mandarin. In *Proceedings of CLS 47*, ed. Carissa Abrego-Collier, Arum Kang, Martina Martinović, and Chieu Nguyen.

Constant, Noah. 2014. Constrastive topic: Meanings and realizations. Doctoral Dissertation, University of Massachusetts.

Dayal, V. 2004. Number marking and (in)definiteness in kind terms. *Linguistics and Philosophy* 27:393–450.

Dayal, Veneeta. 2011. Bare noun phrases. In *Semantics: An International Handbook of Natural Language Meaning*, ed. Claudia Maienborn, Klaus von Heusinger, , and Paul Portner, volume 2, 1088–1109. Mouton de Gruyter.

Dayal, Veneeta. 2012. Bangla classifiers: Mediating between kinds and objects. *Rivista di Linguistica* 24:195–226.

Elbourne, P. 2001. E-type anaphora as NP-deletion. *Natural Language Semantics* 9:241–288.

Elbourne, P. 2005. *Situations and individuals*. Cambridge, MA: MIT Press.

Elbourne, Paul. 2013. *Definite Descriptions*. Oxford: Oxford University Press.

von Fintel, Kai. 1994. Restrictions on quantifier domains. Doctoral Dissertation, MIT.

Fukui, Naoki, and Yuji Takano. 2000. Nominal structure: An extension of the symmetry principle. In *The derivation of VO and OV*, ed. Peter Svenonius, 219–254. Philadelphia: John Benjamins.

Gerner, Matthias, and Walter Bisang. 2010. Classifier declinations in an isolating language: On a rarity in Weining Ahmao. *Language and Linguistics* 11:579–623.

del Gobbo, Francesca. 2003. Appositives at the interface. Doctoral Dissertation, University of California, Irvine.

Greenberg, Joseph. 1978. How does a language acquire gender markers? In *Universals of human language volume 3: Word structure*, ed. Joseph Greenberg, 47–82. Stanford, CA: Stanford University Press.

Groenendijk, Jeroen, and Martin Stokhof. 1990. Dynamic Montague grammar. In *Proceedings of the second symposion on logic and language*, ed. L. Kalman and L. Polos, 3–48. Eotvos Lorand University Press.

Groenendijk, Jeroen, and Martin Stokhof. 1991. Dynamic predicate logic. *Linguistics and Philosophy* 14:39–100.

Hankamer, Jorge, and Ivan Sag. 1976. Deep and surface anaphora. *Linguistic Inquiry* 7:391–428.

Hawkins, John A. 1978. *Definiteness and Indefiniteness: A Study in Reference and Grammaticality Prediction*. London: Croon Helm.

Hawkins, John A. 1991. On (in)definite articles: Implicatures and (un)grammaticality prediction. *Journal of Linguistics* 27:405–442.

Heim, Irene. 1982. The Semantics of Definite and Indefinite Noun Phrases. Doctoral Dissertation, University of Massachusetts, Amherst, PhD.

Heim, Irene. 1990. E-type pronouns and donkey anaphora. *Linguistics and Philosophy* 13:137–177.

Heim, Irene. 1991. Artikel und definitheit. In *Semantics: An international handbook of contemporary research*. Berlin: de Gruyter.

von Heusinger, Klaus. 2013. The salience theory of definiteness. In *Perspectives on linguistic pragmatics*, ed. A. Capone, F. Lo Piparo, and M. Carapezza, 349–374. Springer.

Huang, C.-T. James. 1984. On the distribution and reference of empty pronominals. *Linguistic Inquiry* 15:531–574.

Huang, C.-T. James, Y.-H. Audrey Li, and Yafei Li. 2009. *The syntax of Chinese*. Cambridge, U.K.: Cambridge Unviersity Press.

Ingason, Anton Karl. 2016. Realizing morphemes in the Icelandic noun phrase. Doctoral Dissertation, University of Pennsylvania.

Jenks, Peter. 2015. Two kinds of definites in numeral classifier languages. In *Proceedings of SALT 25*, 103–124.

Jiang, Li. 2012. Nominal Arguments and Language Variation. Doctoral Dissertation, Harvard University, PhD.

Kadmon, Nirit. 1990. Uniqueness. *Linguistics and Philosophy* 13:273–324.

Kamp, Hans. 1981. A theory of truth and semantic representation. In *Formal Methods in the Study of Language*, ed. J. A. G. Groenendijk, T. M. V. Janssen, and M. J. B. Stokhof. Amsterdam: Mathematical Centre.

Kamp, Hans, and Uwe Reyle. 1993. *From discourse to logic*. Dordrecht: Kluwer.

Kang, Arum. 2015. (in)definiteness, disjunction and anti-specificity in korean: A study in the semantics-pragmatics interface. Doctoral Dissertation, University of Chicago.

Karttunen, Lauri. 1969. Problems of reference in syntax. Doctoral Dissertation, Indiana University, Bloomington, IN.

Karttunen, Lauri. 1976. Discourse referents. In *Notes from the linguistic underground*, ed. J. D. McCawley, volume 7 of *Syntax and Semantics*. New York: Academic Press.

Kim, Chonghyuck. 2005. The Korean Plural Marker *tul* and its Implications. Doctoral Dissertation, University of Delaware, PhD.

Kim, Susie, and Hyung-Jo Yoon. 2016. Stop using *That*: Expressing definiteness in korean. *MSU Working Papers in SLS* 7:44–56.

Kratzer, Angelika. 1989. An investigation of the lumps of thought. *Linguistics and Philosophy* 12:607–653.

Kratzer, Angelika. 2007. Situations in natural language semantics. In *Stanford Encyclopedia of Philosophy*, ed. E. N. Zalta. CSLI.

Krifka, Manfred. 1995. Common nouns: A contrastive analysis of Chinese and English. In *The Generic Book*, ed. Gregory N. Carlson and Francis Jeffry Pelletier, 393–411. Chicago, IL: The University of Chicago Press, Chicago.

Kurafuji, Takeo. 2004. Plural morphemes, definiteness, and the notion of semantic parameter. *Language and Linguistics* 5:211–242.

Lee, Chungmin. 1995. Definiteness and specificity. *Linguistics in the Morning Calm (Selected papers from SICOL-1992)* 3:663–677.

Lekakou, Marika, and Kriszta Szendroi. 2007. Eliding the noun in close apposition, or Greek polydefinites revisited. In *UCL working papers in linguistics*, ed. Richard Breheny and Nikolao Velegrakis. University of London, University College London.

Li, C. N., and S. A. Thompson. 1981. *Mandarin chinese: A functional reference grammar*. Berkeley: University of California Press.

Li, XuPing. 2013. *Numeral Classifiers in Chinese: The Syntax-Semantics Interface*. Berlin: Walter de Gruyter.

Li, XuPing, and Walter Bisang. 2012. Classifiers in Sinitic languages: From individuation to definiteness-marking. *Lingua* 122:335–355.

Liao, Wei-Wen Roger, and Yuyun Iris Wang. 2011. Multiple-classifier constructions and nominal expressions in Chinese. *Journal of East Asian Linguistics* 20:145–168.

Lin, Jo-Wang. 2003. On restrictive and non-restrictive relative clauses in Mandarin Chinese. *The Tsing Hua Journal of Chinese Studies, New Series* 33:199–240.

Link, G. 1983. The logical analysis of plural and mass nouns: A lattice theoretic approach. In *Meaning, use, and interpretation of language*, ed. R. Bäuerle, C. Schwarze, and A. von Stechow. Berlin: de Gruyter.

Löbel, Elisabeth. 2000. Classifiers versus genders and noun classes: A case study in Vietnamese. In *Gender in grammar and cognition I: Approaches to gender*, ed. B. Unterbeck and M. Riassanen, volume 124 of *Trends in Linguistics, Studies and Monographs*, 259–319. Berlin: Mouton de Gruyter.

Löbner, Sebastian. 1985. Definites. *Journal of Semantics* 4:279–326.

Löbner, Sebastian. 2011. Concept types and determination. *Journal of Semantics* 28:279–333.

Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.

Merchant, Jason. 2001. *The syntax of silence*. Oxford: Oxford University Press.

Neale, Stephen. 1990. *Descriptions*. MIT Press.

Nguyen, T. H. 2004. The structure of the Vietnamese noun phrase. Doctoral Dissertation, Boston University.

Nomoto, Hiroki. 2013. Number in classifier languages. Doctoral Dissertation, University of Minnesota.

Nowak, Ethan. 2014. Demonstratives, hidden arguments, and presupposition. UC Berkeley.

Oshima, David Y., and Eric McCready. 2016. Anaphoric demonstratives and mutual knowledge anaphoric demonstratives and mutual knowledge. *Natural Language & Linguistic Theory* 1–37. doi:10.1007/s11049–016–9356–6.

Partee, B. 1986. Ambiguous pseudoclefts with unambiguous *Be*. In *Proceedings of NELS XVI*, ed. S. Berman, J.-W. Choe, and J. McDonough. Amherst, MA: University of Massachusetts at Amherst.

Partee, B. 1987. Noun phrase interpretation and type-shifting principles. In *Studies in discourse representation theory and the theory of generalized quantifiers*, ed. J. Groenendijk et al. Foris, Dordrecht.

Patel-Grosz, Pritty, and Patrick G. Grosz. 2017. Revisiting pronominal typology. *Linguistic Inquiry* 48:259–297.

Piriyawiboon, Nattaya. 2010. Classifiers and Determiner-less Languages: The Case of Thai. Doctoral Dissertation, University of Toronto, PhD.

Prince, Ellen. 1998. On the limits of syntax, with reference to topicalization and left-dislocation. In *Syntax and semantics, vol. 29*, 281–302. New York: Academic Press.

Prince, Ellen F. 1981. Toward a taxonomy of given/new information. In *Radical Pragmatics*, ed. Peter Cole. Academic Press.

Roberts, Craige. 1996. Information structure in discourse: Towards an integrated formal theor of pragmatics. In *Papers in semantics*, ed. J. H. Yoon and A. Kathol, volume 49 of *OSU Working Papers in Linguistics*, 91–136. Ohio State University.

Roberts, Craige. 2003. Uniqueness in definite noun phrases. *Linguistics and Philosophy* 26:287–350.

Roberts, Craige. 2004. Pronouns as definites. In *Descriptions and Beyond*, ed. Marga Reimer and Anne Bezuidenhout, 503–543. Oxford: Oxford University Press.

Rullman, Hotze, and Aili You. 2006. General number and the semantics and pragmatics of indefinite bare nouns in Mandarin Chinese. In *Where semantics meets pragmatics*, ed. Klaus von Heusinger and Ken P. Turner, 175–196. Amsterdam: Elsevier.

Russell, Bertrand. 1905. On denoting. *Mind* 14:479–493.

Schlenker, Philippe. 2011. Donkey anaphora: the view from sign language (ASL and LSF). *Linguistics and Philosophy* 34:341–395.

Schlenker, Philippe. 2012. *Maximize Presupposition* and Gricean reasoning. *Natural Language Semantics* 20:391–429.

Schwartzschild, Roger. 1996. *Pluralities*. Dordrecht: Kluwer Academic Publishers.

Schwarz, Florian. 2009. Two Types of Definites in Natural Language. Doctoral Dissertation, University of Massachusetts, Amherst, PhD.

Schwarz, Florian. 2012. Situation pronouns in Determiner Phrases. *Natural Language Semantics* 20:431–475.

Schwarz, Florian. 2013. Two kinds of definites cross-linguistically. *Language and Linguistics Compass* 7:534–559.

Simpson, A. 2005. Classifiers and DP structure in Southeast Asia. In *The Oxford Handbook of Comparative Syntax*, ed. Guglielmo Cinque and Richard Kayne, 806–838. Oxford University Press.

Simpson, Andrew. 2017. Bare classifier/noun alternations in the Jinyun (Wu) variety of Chinese and the encoding of definiteness. *Linguistics* 55:305–331.

Simpson, Andrew, and Priyanka Biswas. 2016. Bare nominals, classifiers, and the representation of definiteness in Bangla. *Linguistic Analysis* 40:167–198.

Simpson, Andrew, Hooi Ling Soh, and Hiroki Nomoto. 2011. Bare classifiers and definiteness. *Studies in Language* 35:168–193.

Sio, Joanna U.-S. 2006. Modification and reference in the Chinese nominal. Doctoral Dissertation, Leiden University.

Szabolcsi, A. 1994. The noun phrase. In *Syntax and semantics, vol 27: The structure of Hungarian*, ed. Ferene Kiefer and Katalin É. Kiss. London: Academic Press.

Trinh, Tue. 2011. Nominal reference in two classifier languages. In *Sinn und Bedeutung 15*, ed. Ingo Reich, 629–644.

Wolter, Lindsay. 2006. That's that: The semantics and pragmatics of demonstrative noun phrases. Doctoral Dissertation, UC Santa Cruz.

Wu, Yicheng, and Adams Bodomo. 2009. Classifiers ≠ determiners. *Linguistic Inquiry* 40:487–503.

Yang, Rong. 2001. Common Nouns, Classifiers, and Quantification in Chinese. Doctoral Dissertation, Rutgers University, PhD.

Zhang, Niina Ning. 2015. Nominal-internal phrasal movement in Mandarin Chinese. *The Linguistic Review* 32:375–425.

Peter Jenks
Department of Linguistics
UC Berkeley
1203 Dwinelle Hall #2650
Berkeley, CA 94720-2650

jenks@berkeley.edu