The Puzzle of Anaphoric Bare Nouns in Mandarin: A Counterpoint to *Index!*

Veneeta Dayal and Li Julie Jiang

Abstract: Jenks (2018) argues that Mandarin bare NPs cannot be classified as definites simpliciter. Adopting the distinction between weak and strong article definites in Schwarz (2009), he proposes that Mandarin makes a lexical distinction between the two types of definites: bare nouns are weak definites, demonstratives are strong definites. He further proposes that their distribution is regulated by a principle called *Index!*. In this article we first point out some problems with the empirical generalizations presented in Jenks' description of Mandarin and then sketch out an alternative approach to the distinction between Mandarin demonstratives and bare nouns. We end with some comments about the kind of further empirical work that needs to be done before definitive claims can be made about the competition between demonstratives and other types of definites.

Keywords: strong and weak definite articles, bare noun, demonstrative, Mandarin cross-linguistic variation, competition

1 Weak and Strong Definiteness in Mandarin

Mandarin is known to lack an overt definite article, expressing definiteness via bare nouns (e.g. Chao 1968; Li and Thompson 1981; Cheng and Sybesma 1999; Yang 2001, among others) (1a). The ability of bare nouns to function as definites is also illustrated by (1b), in which the linguistic context brings out their anaphoric use (e.g. Dayal 2004, 2011; X. Li 2013; Jiang 2012, 2020):

(1) a. Hufei he-wan-le **tang**.

Hufei drink-finish-LE soup

'Hufei finished the soup.' (Cheng and Sybesma 1999: 510)

b. Wo kanjian yi zhi mao. **Mao** zai huayuan-li.

I see one CL cat. Cat at garden-inside

'I see a cat. The cat is in the garden.' (Dayal 2004: 403)

Jenks (2018) modifies this generalization on the basis of examples like (2). According to him, the anaphoric potential of bare nouns does not extend beyond the subject position. The bare nouns in (2b, c) are infelicitous in both direct object and indirect object positions, requiring the use of demonstratives instead:

(2) a. Jiaoshi li zuo zhe yi ge nansheng yi ge nüsheng.

classroom inside sit PROG one CL boy one CL girl

'There is a boy and a girl sitting in the classroom.'

b. Wo zuotian yudao #(**na ge**) **nansheng**.

I yesterday meet that CL boy

'I met the boy yesterday.'

c. Wo dai gei #(**na ge**) **nansheng** yi ge liwu.

I bring give that CL boy one CL gift

'I'm bringing a gift for the boy.' (Jenks 2018: 510)

Jenks further argues that only uniqueness-based definites, or weak-article definites in the sense of Schwarz (2009), are realized with bare nouns in Mandarin and that demonstratives can occur in these contexts only with contrastive interpretations. He examines three environments that require unique definites, using the classification in Hawkins (1978). In addition to bare nouns functioning as immediate-situation definites, as seen in (1a), they also function as larger-situation definites, demonstrated in (3a), and as associative definites in part-whole bridging contexts shown in (3b) (for more on associative anaphora, see Clark 1975; Hawkins 1978; Schwarz 2009):

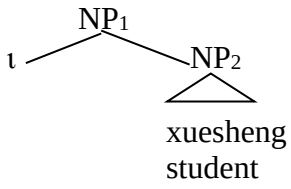
(3) a. **Yueliang** sheng shang lai le.
Moon rise up come LE
'The moon has risen.' (Chen 2004:1165)

b. Chezi bei jingcha lanjie le yinwei mei you tiezhi zai **paizhao** shang.
car PASS police intercept LE because NEG have sticker at license.plate on
'The car was intercepted by the police because there wasn't a sticker on the license plate.' (Jenks 2018: 507-508)

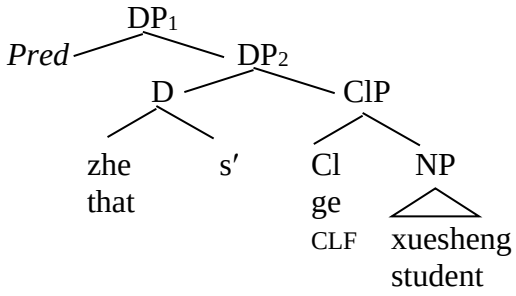
Jenks offers the following analysis to account for the contrast between Mandarin bare nouns and demonstrative expressions. First, he assumes that Mandarin bare nouns project NPs (4a), whereas demonstrative expressions project DPs, consisting of at least three distinct nominal projections: DP>ClfP>NP (4b). Second, he adopts Trinh's (2011) analysis

of common nouns and numeral classifiers and treats Mandarin bare nouns as predicative; he further adopts the analysis that Mandarin bare nouns achieve their definite interpretation via the type-shifting operator *iota* ι (e.g. Yang 2001) (5a).[1] Third, he follows Schwarz's (2009) account of weak vs. strong definites, which are distinguished by the presence of an index only on strong anaphoric definites. Specifically, he proposes that Mandarin demonstratives are like strong-article anaphoric definites in German in taking an extra semantic argument, which in most contexts is satisfied by an index ι^x (5b). He assumes that the index is interpreted as an indexical property relativized to an assignment function λx[x = *g* (1)] and that the indexical argument of ι^x is represented as a DP adjunct:[2]

(4) a. Unique definites



b. Anaphoric definites



(5) a. $[\![NP_1]\!]^g = \exists!x[\textbf{student}(x)(s')].\iota x[\textbf{student}(x)(s')]$

b. $[\![DP_1]\!]^g = \exists!x[\textbf{student}(x)(s') \wedge AT_{obj}(x) \wedge x = g(1)].$

$\iota x[\textbf{student}(x)(s') \wedge AT_{obj}(x) \wedge x = g(1)]$

Jenks follows the literature on bare nouns in interpreting the Mandarin common noun as a kind denoting term. He posits a covert *iota* type shift in the case of bare noun arguments. Details aside, we get the expected meaning for the noun phrase: a presupposition that there is a unique student in the situation and, when defined, the NP refers to that individual. In the case of the demonstrative, a classifier phrase intervenes.[3] Again, as is standard, classifiers take a kind term and deliver a set of atomic or singular individuals (AT_{obj}). To this the demonstrative, qua strong definite, adds the indexical property. Here Jenks departs slightly from Schwarz (2009) but the essential idea is the same. Though *iota* is still implicated, its uniqueness requirement is defined on two properties, the property of being a student and of being identical to a previously introduced individual (x = g(1)). As in Schwarz's original account, this distinction is meant to separate out the contexts in which uniqueness of the common noun is at stake from the contexts in which uniqueness rides on indexicality. The exponent for the first type is the weak article in German and the exponent for the second type is the strong article in German.[4] On Jenks' account the exponent for the first type is the bare noun in Mandarin and the exponent for the second type is the demonstrative in Mandarin.

According to Jenks, then, Mandarin patterns with German in making a principled distinction between weak (unique) and strong (anaphoric) definites. It differs from English which uses a single form for both types of definites. The resulting cross-linguistic picture is given in table 1.

Table 1

Definiteness marking in Mandarin, German, and English

	Mandarin	German	English
Unique definites	N	Weak article	*the*
Anaphoric definites	Dem-Cl-N	Strong article	*the*

(Jenks 2018: 529)

The core proposal in Jenks (2018) is a principle called *Index!*, which requires Mandarin to explicitly represent indices whenever possible:

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

Since the claim is that Mandarin demonstrative expressions include an index that is absent in definite bare nouns, *Index!* would require demonstratives to be used whenever they can. In the case of (2b, c), demonstrative expressions are available since they are anaphoric to expressions in the antecedent clause and therefore must be used; as a result, bare nouns become unavailable in those contexts.

Jenks makes one exception to *Index!* As seen in (1b), subject bare nouns can be anaphorically linked to expressions in antecedent clauses. To account for this, he proposes that anaphoric bare nouns in subject position are continuing topics in Mandarin. The pragmatic function of topic marking overrides and neutralizes the effect of *Index!* in

such environments since topics are salient members of the question under discussion and do not need to be indexed.

This, in brief, is the account in Jenks (2018) which combines the view of Mandarin bare nouns as kind denoting terms with the view of definiteness as being composed of a uniqueness-based sub-type and an anaphora-based subtype.

2 Fault Lines in the Empirical Foundations of *Index!*

The data presented in Jenks (2018) certainly changes the picture of Mandarin bare nouns that was earlier assumed in the literature. However, the conclusions that he would have us draw are not quite supported when a larger set of facts are taken into account. In this section we take a closer look at the two empirical generalizations that form the basis of Jenks' proposal: the behavior of demonstratives in contexts that favor weak article definites and the behavior of bare nouns in contexts that favor strong article definites. In each case, we fill in some crucial gaps in the paradigm in order to lay a more solid foundation for claims about the lexical exponents of definiteness in Mandarin.[5]

2.1. Mandarin demonstratives as ordinary demonstratives

Schwarz (2009, 2013) bases his division of definites on languages that show a lexical distinction in behavior, languages like German, Lakhota, and Akan, and argues that the distinction between strong and weak definiteness holds cross linguistically.[6] There are languages like English, however, where this distinction is not apparent but on Schwarz's

view this is because a single definite article (*the* in English) is ambiguous in encoding both weak and strong definiteness.

Schwarz is careful to note that strong and weak definite articles do not exhaust the range of possibilities for the expression of definiteness and mentions, among other expressions, demonstratives as an example (see in particular Schwarz 2009: 34-37). Since on Jenks' account the Mandarin demonstrative *is* the strong definite, we believe it is worth understanding what, if anything, differentiates a strong definite from a demonstrative. We can do so with reference to English where the two are lexically distinct.

The larger situation use and part-whole bridging contexts mentioned in section 1 provide a good starting point. In both, the set denoted by the common noun head is uniquely instantiated – world knowledge tells us that there is only one moon that we ordinarily talk of and only one license plate that a given car can have. While the English definite is acceptable in such cases, the English demonstrative is not:

(7) a. The/#That moon has risen.

b. The police stopped the car because the/#that license plate was not visible.

c. DEF_{WEAK} / #$\text{DEF}_{\text{STRONG}}$ / #DEM N

On the ambiguity view of English definites, we can analyze the situation as in (7c). The weak definite survives, the strong definite and the demonstrative are ruled out. Note that the infelicity of Mandarin demonstratives in the corresponding cases seen in (3) is

equally compatible with their being analyzed as plain old demonstratives, as with their being analyzed as strong definites. These contexts therefore provide no reason to move from the null hypothesis, which is that Mandarin demonstratives are demonstratives rather than definites.

There is, however, one further case discussed by Jenks that is meant to settle the question in favor of Mandarin demonstratives being ambiguous between demonstratives and strong definites. These are associative nouns that do not involve part-whole relationships of the kind seen in (3b) and (7b). Schwarz notes that several languages differentiate between these two types of associative uses and while prohibiting strong definites from the part-whole cases, allow them in producer-product type cases.[7] Jenks provides the example in (8a) to establish that Mandarin allows the demonstrative to be used here, aligning it with languages that allow strong definites in these cases:

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

b. Paul du-le yi ben youqu-de shu. Ta xiang jian **zuozhe**.
Paul read-PERF one CLF interesting-MOD book. he want meet author.
'Paul read an interesting book. He wants to meet the author.'

The data in (8a) is meant to settle the case in favor of Mandarin demonstratives being strong definites since the English translation of the associated NP *poet* is acceptable with the definite determiner (and presumably not with a demonstrative). The explanation is as follows. The ordinary demonstrative meaning of *na wei shiren* 'that Clf poet' is ruled out analogously to the way *that poet* would be ruled out in English, but its strong definite meaning is acceptable analogously to the way *the poet* is acceptable. We believe this conclusion is too hasty.

Note that the antecedent noun in the main clause of (8a) also has a demonstrative *na shou shi* 'that Clf poem'. The canonical examples of bridging in the literature typically have indefinite antecedents. When we adjusted the data to have an overt indefinite antecedent, as in (8b), the judgements changed and our consultants readily accepted the bare NP.[8] We therefore do not see any basis for making the distinction between part-whole and producer-product sub-cases of associative anaphoric nouns in Mandarin. Consequently, we do not see sufficient empirical grounds to consider the Mandarin demonstrative anything other than what it appears to be, a regular demonstrative.

There is also a theoretical issue with the way *Index!* relates to strong definites that is worth flagging. Consider the sequence of English sentences in (9a-b):

(9) a. The President was talking to a minister. The minister was asking the/#that President …

b. The sun and the moon are part of our solar system. The earth revolves around the/#that sun…

c. DEF$_{WEAK}$……..DEF$_{WEAK/STRONG}$…..

The first mention of *the President/the sun* involves a weak definite, an instance of larger situation uniqueness. It is not so clear how the second instance of these noun phrases should be analyzed. They still satisfy uniqueness, so we could use the weak definite but they also satisfy anaphoricity, so *index!* might tell us to use the strong definite. Again, English does not help us in this regard because the ambiguity of *the* masks the distinction, but the Mandarin anaphoric noun is transparent:

(10) a. Zongtong zhengzai gen yi ge buzhang shuohua.
president PROG with one CL minister talk
Buzhang wen **(#na ge) zongtong**…
minister ask that CL president
'The President was talking to a minister. The minister was asking the/#that President…'

b. Taiyang he yueliang shi women taiyangxi de yi bufen.
sun and moon be our solar.system MOD one part.
Diqiu weirao **(#na ge) taiyang** zhuan…
earth revolve that CL sun turn
'The sun and the moon are part of our solar system. The earth revolves around the/#that sun…'

c. DEF$_{WEAK}$……..DEF$_{WEAK/STRONG}$…..

Both instances of *president/sun* have to be bare, demonstratives in the second sentence are infelicitous. So, as in the other cases discussed here, the Mandarin demonstrative turns out to pattern with the English demonstrative, not with the English strong definite. We should point out that Jenks in a sense inherits this problem from Schwarz but by putting a principle like *Index!* in place, he makes a specific prediction for Mandarin that is not borne out.[9] We will come back to the appropriate account of demonstratives in section 3. We turn now to the other side of the equation and test the generalization that non-subject Mandarin bare nouns are infelicitous in anaphoric contexts.

2.2 Mandarin Bare NPs as Anaphoric Definites

Recall from (2) that anaphoric bare nouns in direct object and indirect object positions were claimed to be infelicitous in Jenks (2018). We observe, however, that this is not always the case. For example, when the second sentence contains two bare nouns, one in subject position and one in object position, they can both refer anaphorically to the indefinite expressions in the first sentence:

(11) a. Jiaoshi li zuo zhe yi ge nansheng yi ge nüsheng.
classroom inside sit PROG one CL boy one CL girl
'There is a boy and a girl sitting in the classroom.'

b. **Nüsheng** zuo zai **nansheng** pangbian.

girl sit DUR boy side

'The girl was sitting next to the boy.'

c. **Nüsheng** zhengzai gen **nansheng** shuohua.

girl PROG with boy talk

'The girl was talking to the boy.'

Note that the bare nouns in (11b, c) are not subject to the exception Jenks makes for bare nouns in subject position. Crucially, while demonstratives are judged to be possible in subject and object positions in these examples, all the native speakers we consulted found two bare nouns to be acceptable, several even voicing a preference for two bare nouns over a demonstrative in one or both positions.

It has been suggested to us that having two bare nouns in the second sentence may create a contrastive environment and that *Index!* can be overridden and neutralized by contrastive focus, just as it can by continuing topics (Peter Jenks, pers. comm.). If focus may also introduce an index of its own, the occurrence of bare nouns in object position in (11) is to be expected under *Index!*. Although this possibility can help maintain the proposal in Jenks (2018), we do not find it compelling, theoretically or empirically. Naturally occurring data establish that Mandarin bare nouns can be anaphoric in non-subject positions in sentences which are not contrastive in any identifiable way. Below, we provide some examples from the Beijing Language and Culture University Contemporary Chinese Corpus (BCC Corpus in short):[10]

(12) a. ‘Shenchu shou lai, bishang yanjing.’ Yi ming huanyou jingshen zhang’ai
extend hand come close eye one CL have mental disorder
de nanhai mingling Youyou. Youyou shuncongde ba shou di-gei **nanhai**.
DE boy order Youyou Youyou obediently BA hand hand-to boy
‘Show me your hand and close your eyes.’ One boy who has mental disorder ordered Youyou. Youyou showed his hands to the boy obediently.’

(BCC Corpus, from China Daily, Dec 7, 2016)

b. Yi tiao gou dagai kan chuan chang-yi lache de bu shen
one CL dog probably see wear long-clothes pull-rickshaw DE not very
shunyan, gen zhe ta yao. Ta tingzhu le che,
pleasing.to.the.eye follow PROG 3SG bite 3SG stop PERF rickshaw,
dao zhuai zhe buzi, pinming de zhui zhe **gou** da.
opposite grab PROG dusk.whisk try.very.hard DE race PROG dog beat
‘One dog probably didn’t find the rickshaw puller in the long shirt pleasing to the eye and was following him to bite him. He stopped the rickshaw, grabbed his dust-whisk by the whisk-end and raced very hard after the dog.’

(BCC Corpus, from *Rickshaw Boy*, by Lao She)

The bare nouns, *nanhai* ‘boy’ and *gou* ‘dog’, appear as objects in the second clause, referring anaphorically to indefinite expressions in the antecedent clauses. Note that the

subjects in the second clause in (12a) and (12b) are a proper name and a pronoun respectively.

The example in (13) further illustrates that bare nouns in indirect object position can also refer anaphorically to the indefinite in the antecedent clause:

(13) Yi waiguo nühai kandao lubian mai de xiaogou, hen xihuan dan mei qian
one foreign girl saw street-side sell DE puppy very like but not money
mai, jiujiu bu ken liqu, yushi mailai shui, bian wei gougou bian luo lei.
buy long not willing leave then buy water while feed dog while shed tear
Mai gou dashu zuizhong ba liang zhi gougou song gei le **nühai**.
sell dog uncle finally BA two CL puppy give to PERF girl
'One foreign girl saw puppies being sold on the street, (she) liked them very much but had no money to buy and didn't want to leave; then (she) bought water, shedding tears while feeding the puppies. In the end, the man who was selling puppies gave two puppies to the girl.' (BCC Corpus, from Weibo 'Microblog')

It is worth emphasizing that in the examples in (12) and (13), demonstrative expressions can also appear where the anaphoric bare nouns appear. *Index!* incorrectly predicts a competition between these two forms, with bare nouns losing out to demonstratives.

To make this point further, (14) illustrates that anaphoric bare nouns can appear multiple times in object (and subject) position in the second clause and clauses thereafter:

(14) a. Ou Weiling turan faxian yi ge zhongnian funü zai douyin
Ou Weiling suddenly notice one CL middle.aged woman PROG teasing
yi ge nühai, xiang jiao **nühai** gen ta hui jia qu, dan **nühai**
one CL girl want ask girl with 3SG return home go but girl
wensibudong. Ou Weiling juede youdian chayi, bian shangqian
not.move.a.single.jot Ou Weiling feel a.little weird then go.forward
wen **nühai**: 'ni zai deng shenme ren?'
ask girl you PROG wait which person
'Ou Weiling suddenly noticed one middle-aged woman was teasing one girl, (she) wanted to ask the girl to go back home with her, but the girl didn't move a single jot. Ou Weiling felt a bit weird and hence stepped forward to ask the girl: 'who are you waiting for?'

(BCC Corpus, from Xiamen Daily, Dec 13, 1993)

b. Zai shuihu bangbian yi zhi cuxia-xiang de mao zai canzhuo shang
At kettle side one CL mischievous-look DE cat on desk top
shuizhao le. Laozuoli'en like ba **mao** xu-zou.
sleep SFP Old-Jolyon immediately BA cat drive-away
Ta ba da limao pai-de pa-pa zuoxiang, yi qugan zhe **mao**.
3sg BA big hat slap-de bang-bang sound to chase-away PROG cat
'Besides the kettle, one mischievous-looking cat fell asleep on the desk. Old Jolyon immediately drove the cat away. He slapped the big hat very loud in order to chase away the cat.

(BCC Corpus, from *The Man of Property*, by John Galsworthy)

In (14a), the anaphoric bare noun *nühai* ‘girl’ first appears in object position in the second clause and then appears in subject position in the third clause and object position in the fourth clause. In (14b), the anaphoric bare noun *mao* ‘cat’ first appears in the *ba-construction* in the second clause, leading to the SOV (as opposed to SVO) word order, and appears again in object position in the third clause.

We have provided evidence based on elicitation from six native speakers of Mandarin and from naturally occurring sentences from a major corpus against the empirical generalizations in Jenks (2018). We have more such examples which for reasons of space we do not include here. This suggests to us that *Index!* cannot explain the distribution of demonstratives and bare nouns in Mandarin and an alternative is worth exploring. We show next what an account with the potential to capture the range of facts that we now have in front of us might look like.

3 An Alternative to *Index!*

In this section we suggest a way of analyzing Mandarin demonstratives that aligns them with demonstratives cross-linguistically, while delivering the apparent strong definite behavior noted in Jenks (2018). We also speculate on an alternative way of capturing the crucial contrasts in the anaphoric behavior of Mandarin bare nouns, that is, between the examples on which Jenks bases his claims for their non-anaphoricity and the examples we have provided to show their anaphoricity.

3.1. The Anti-Uniqueness of Demonstratives

Let us step back a bit from the specifics of Mandarin and ask ourselves what we know about demonstratives generally. One observation in the literature is that most, perhaps all, languages have demonstratives, while a large number of languages lack definite articles. Another is that demonstratives are the historical source for definite articles. There is a rich literature on the semantics of demonstratives, of course, but we will take our inspiration from Robinson (2005) as it bears most closely on the issue of interest here (see also Roberts 2002, Wolter 2006). Drawing on Löbner (1985), Robinson notes that demonstratives differ from definites in their ability to tolerate lack of uniqueness:

(15) a. This dog is awake and this dog is asleep.

b. #The dog is awake and the dog is asleep.

In fact, Robinson argues, a demonstrative not only tolerates non-uniqueness, it requires it:

(16) a. #That sun is hot.

b. Helen bought a car. #That steering wheel is dangerous.

c. The match was interesting. #That umpire was unfair.

Given these facts, she suggests that although demonstratives refer to a unique entity, they have a presupposition of non-uniqueness. We can represent it as in (17), and given that these facts hold cross-linguistically we can take (17) to also hold in Mandarin:

(17) $[\![\text{Dem}]\!] = \lambda P: |P| > 1.\ \iota x\ [P(x) \wedge x = y]$ *initial proposal*

That is, each instance of a demonstrative picks out a unique entity that satisfies two properties, the common noun property and the intended referent property, here represented by the free variable *y,* in line with the convention in Schwarz (2009). The crucial difference is in the presuppositional piece of the demonstrative's meaning. It requires that the intended referent not be the sole member of the common noun set. It differs from the definite where there is a presupposition of uniqueness on the common noun set: $|P| = 1$.

Let us illustrate how this explains the contrasts in (15) and (16). Since $|\text{dog}| = 2$ in (15), each conjunct satisfies the non-uniqueness presupposition as well as the entailment of uniqueness – there are two dogs in the context but the demonstratives in the two conjuncts have distinct indices, ensuring uniqueness calibrated to those indices. The problem with (16) is with the presupposition of non-uniqueness. For (16a) normal world knowledge tells us that there is only one sun in our conceptual universe and for (16b-c), that there is only one *steering wheel/umpire* per car or match.[11]

Based on what we have said so far, one may expect demonstratives and definites to be in complementary distribution, but we know that is not the case. Most relevant to our concerns are anaphoric contexts where both options seem possible in English:

(18) a. A woman and a man came into the room. The woman sat down.

b. A woman and a man came into the room. That woman sat down.

Robinson notes, and we agree, that the version with the definite seems unmarked, while the version with the demonstrative suggests a slight sense of contrast. Setting that aside, how can we explain the difference between cases like (18b) and (16), given the proposed semantics in (17)? In order to do so, we minimally modify the non-uniqueness presupposition in (17), as in (19a), to allow its satisfaction in a larger situation:[12]

(19) a. $[\![\text{Dem}]\!] = \lambda s\, \lambda P\colon \exists s'\ s \leqslant s'\ |P_{s'}| > 1.\ \iota x\ [P_s(x) \wedge x = y]$ *final proposal*

b. $[\![\text{the}_{\text{STRONG}}]\!] = \lambda s\, \lambda P\colon |P_s \cap \lambda x[x=y]| = 1.\ \iota x\ [P_s(x) \wedge x = y]$ *strong definite*

c. $[\![\text{the}_{\text{WEAK}}]\!] = \lambda s\, \lambda P\colon |P_s| = 1.\ \iota x\ [P_s(x)]$ *weak definite*

We are importing the notion of widening, proposed by Kadmon and Landman (1993) to explain the polarity item *any*, for the satisfaction of the non-uniqueness condition in (19a). Domain widening has also been used by Dayal (2013) to explain the indefiniteness typically associated with bare plurals in episodic contexts. On the opposite side of the spectrum, D-linked expressions like *each* are thought to resist such widening (Kadmon

and Landman 1993: 378- 379, Dayal 2016: 122-124). Here we suggest that the initial situation invoked by the first sentence in (18b), with a unique woman and a unique man, is extended to include a larger situation which opens up the possibility of having other women in it, even if no such woman is salient. And this is what licenses the demonstrative in the second sentence of (18b). The strong definite in (18a) is also felicitous since its uniqueness presupposition is satisfied in the situation in which the first sentence is interpreted. For completeness we include the weak definite. In this situation it also happens to be satisfied since there is not only a unique element in the intersection of P and the set of individuals anchored to the indefinite in the first sentence, but in P itself. To anticipate, we propose that the Mandarin demonstrative is to be represented as (19a), and the Mandarin bare noun in terms of (19b) and (19c).[13] The choice between the strong and the weak definite is not trivial and we will return to it in the following sections but let us keep our focus on the demonstrative for now.

Our account of demonstratives rests on the view that domain widening in anaphoric cases such as (18b) involves a relatively smooth transition from a more restricted situation to a larger situation. One might ask why it is not similarly possible to license a demonstrative on a widened domain in larger situation uses of the kind illustrated in (16a). The answer seems to be the following. To allow the possibility of other suns, we would need to adjust our context to incorporate other solar systems. While such adjustments are certainly possible, it seems that they require some rather serious adjustments to our initial assumptions. And in the case of associative readings such as the ones we show in (16b-c), the demonstrative remains anchored to a single entity through a

relationship defined on uniqueness. No matter how many other cars or matches there may be in the widened domain, there will still be only one steering wheel and one umpire anchored to the car or match referred to in the first sentence. To jump to the steering wheel of other possible cars or umpires of other possible matches seems to be a bridge too far for the average speaker.[14] However, note that immediate situation uses involving ordinary common nouns, as in *that boy is tall*, would be acceptable under what would be described as a deictic use of the demonstrative.

Let us take stock of where we are. We have drawn on Robinson's insight about the difference between definites and demonstratives to argue for the role of a non-uniqueness condition on the use of demonstratives and suggested a way in which their distribution can be differentiated from that of the definite determiner. Given that the behavior of demonstratives in core cases is relatively stable across English and Mandarin (and, as far as we know, most languages), we have suggested that Mandarin demonstratives are really just regular demonstratives. This, one might even argue, is the null hypothesis for, as we demonstrated in section 2.1, there is no empirical imperative to treat Mandarin demonstratives as ambiguous between demonstratives and strong definites.

3.2. Anaphoric Bare Nouns, Situations and Individuals

While we presented ample evidence in section 2.2 against the generality of Jenks' (2018) conclusion that Mandarin bare nouns cannot be anaphoric definites, we do not challenge

his core data. We repeat a near minimal pair to remind us of the contrast at issue and to highlight the nature of the puzzle posed by anaphoric bare nouns in Mandarin:

(20) a. Jiaoshi li zuo zhe yi ge nansheng yi ge nüsheng.
classroom inside sit PROG one CL boy one CL girl
'There is a boy and a girl sitting in the classroom.'

b. **Nüsheng** zuo zai **nansheng** pangbian.
girl sit DUR boy side
'The girl was sitting next to the boy.' ((11b) in section 2.2)

c. Wo zuotian yudao #(**na ge**) **nansheng**.
I yesterday meet that CL boy
'I met the boy yesterday.' ((2b), from Jenks 2018: 510)

There are two things we can observe about these examples. The first is that the syntactic position does not determine the shape of the anaphoric definite – there are non-subject anaphoric nouns, *nansheng* 'boy', in both sentences and yet one is a bare noun while the other is a demonstrative. The second observation is that we are talking about preferences, not judgments of absolute (un)grammaticality. There is a preference for a demonstrative in cases like (20c), but the bare noun is not exactly ruled out for all speakers. The bare noun is fully acceptable for all speakers in (20b), but the demonstrative would not be deemed ungrammatical. One conclusion we can safely draw is that *Index!* does not provide the right level of granularity in the explanation to capture the nuances of this

paradigm. The challenge then is to find a different angle from which to approach the problem. Here we will build on a suggestion made by Gita Martohardjono (pers. comm.) to see if it can shed light on the observed contrast.

Let us start by asking how the context setting sentence in (20a) would be interpreted. On a situation semantic approach, we might posit something like (21a), taking *s* to be the minimal situation which has a boy and a girl in it:

(21) a. ∃s ∃x ∃y [boy(x, s) ∧ girl(y,s) ∧ in-classroom(x,s) ∧ in-classroom(y,s)]

b. ∃!y[girl(y,s')] ∧ ∃!x[boy(x,s')].

∃s' [sitting-next-to (ιy[girl(y,s')], ιx[boy(x,s')]) (s')]

c. ∃s" s'⩽s" $|\text{boy}_{s"}| > 1$.

∃s' ∃u [speaker(u, s') ∧ met-yesterday (u, ιz[boy(z,s') ∧ z=x]) (s')]

Now let us consider the two follow-up sentences in (20b) and (20c), in isolation for now. In (20b) there are two bare nouns, *nüsheng* 'girl' and *nansheng* 'boy', each of which can be considered to be an instance of an immediate situation unique definite. This is derivable via the *iota* type shift (Chierchia 1998, Cheng and Sybesma 1999, Yang 2001, Dayal 2004, Jiang 2012, 2020, Jenks 2018, among others). For perspicuity we include the presupposition on a separate line at the top. In (21b) we have the presupposition that there is a unique boy and a unique girl in the situation of evaluation and the assertion that the unique girl is sitting next to the unique boy. Interpreted in isolation, (20b) is an instance of immediate situation definites in Hawkins' classification and is covered by weak article

definites in Schwarz (2009). Note that this is how the bare noun *tang* 'soup' in example (1a) is interpreted as a definite, on our account as in Jenks'. It also covers larger situation uses of bare nouns like *yueliang* 'moon' in example (3a), where the situation of evaluation *s'* would be identified with the world w_s.

In (20c), we have instead of a bare noun, *na-ge-nansheng* 'Dem-Cl-noun'. The truth conditional contribution of the demonstrative is the same as that of a strong definite, but on the account we sketched out in section 3.1 it has a presupposition of non-uniqueness on the set of boys, satisfiable on a widened domain. Again, taking (20c) by itself, we predict it to be acceptable. And, indeed, it is. In immediate situation uses it would translate into a deictic use of the demonstrative ('I met that boy yesterday.'), a use that Jenks recognizes, as would any account of demonstratives. Our account also has the advantage of indicating why demonstratives cannot function as larger situation definites: *yueliang* 'that-Cl-moon' will not satisfy the non-uniqueness presupposition of demonstratives as it is unique in the widest situation available, the world of the situation w_s. Note that the unacceptability of the demonstrative in larger situation uses (e.g., (3a)) is categorical, not just a preference for the bare noun. This is expected since the unacceptability is due to presupposition failure.

With the basics in place, we return to anaphoric contexts. The contrast between (20b) and (20c) emerges when they are seen as continuations of (20a) in a narrative sequence. Let us illustrate what happens when (21b-c) are treated as follow-ups to (21a). We get the following:

(22) a. ∃s ∃x ∃y [boy(x, s) ∧ girl(y,s) ∧ in-classroom(x,s) ∧ in-classroom(y,s)

∧ **∃s' s ≤ s'** [sitting-next-to (ιy[girl(y,s')], **ιz[boy(z,s')**]) (s')]]

b. ∃s ∃x ∃y [boy(x, s) ∧ girl(y,s) ∧ in-classroom(x,s) ∧ in-classroom(y,s)

∧ **∃s' s ≤ s'** ∃u [Speaker(u, s') ∧ met-yesterday (u, **ιz[boy(z,s') ∧ z=x]**) (s')]]

As the logical representations make explicit, both the bare noun and the demonstrative are defined in this context and end up referring to the same individual. That is, our account shows why both discourses are possible but we need to identify what could explain the preference for a bare noun in the first case and for a demonstrative in the second. The key, we believe, is in the relation between the initial situation *s* and the subsequent situation *s'*. In the version of situation semantics in which Schwarz couches his account of strong and weak article definites there is no ontological difference between situations and individuals (Schwarz 2009: 223). Given that, there is nothing substantive to the condition $s \leq s'$ in (22a) but there is one in (22b): s is defined on two individuals and the same two individuals define s' in (22a) but the individuals in s are a proper part of the individuals in s' in (22b). We can therefore rewrite the formulas as in (23):

(23) a. ∃s ∃x ∃y [boy(x, s) ∧ girl(y,s) ∧ in-classroom(x,s) ∧ in-classroom(y,s)

∧ sitting-next-to (ιy[girl(y,s)], **ιz[boy(z,s)**]) (s)]

b. ∃s ∃x ∃y [boy(x, s) ∧ girl(y,s) ∧ in-classroom(x,s) ∧ in-classroom(y,s)

∧ **∃s' s ≤ s'** ∃u [Speaker(u, s') ∧ met-yesterday (u, **ιz[boy(z,s) ∧ z=x]**) (s')]]

One might argue that since the initial situation extends across the whole discourse in (23a), the simplest type of definite would be one that employs the simple type-shift *iota,* which in the case of Mandarin is the bare noun. And, given that there are two distinct situations in (23b), a deictic definite might have an advantage. While there is no reason to doubt that there is still only the unique boy in *s'*, we can see that there are now potentially two situations in play, a situation *s* defined by a boy and a girl, and a situation *s'* that presumably expands the original situation to include another individual. Though there is no indication that there are other boys in *s'*, the demonstrative can refer unambiguously to the unique boy in the initial situation *s*. We might represent this schematically, as in (24). While, we have linguistic evidence only for the distinction in (24a) and (24b), we speculate that once the possibility of a distinct s' comes into the picture, it brings in the possibility of extensions such as (24c) and that is an extension in which *iota* (the operation that the definite bare noun taps into) is infelicitous but the demonstrative remains felicitous:

(24)	a. {boy$_1$, girl$_1$}	⇒	{boy$_1$, girl$_1$}	*iota: felicitous*
	s		s'	*demonstrative: felicitous*
	b. {boy$_1$, girl$_1$}	⇒	{boy$_1$, girl$_1$, speaker}	*iota: felicitous*
	s		s'	*demonstrative: felicitous*
	c. {boy$_1$, girl$_1$}	⇒	{boy$_1$, girl$_1$, speaker, boy$_2$}	*iota: infelicitous*
	s		s'	*demonstrative: felicitous*

Looking at the contrast in terms of situations and the entities that constitute them provides us with a fresh angle from which to address this puzzle. We no longer predict complementary distribution but we do allow for the possibility of a preference. If speakers feel confident that the initial situation remains unchanged, they have a choice between two felicitous options and they choose the simpler option, namely the bare noun encoding the simple type-shift *iota*. Once the original situation is extended, however, speakers may play it safe and choose the demonstrative which would remain felicitous even if the extension was drastic, over the definite which could become infelicitous.

If this line of thinking has any merit, it predicts that anaphoric bare nouns should improve in the following, where the first sentence itself sets up a minimal situation with all three individuals, a boy, a girl and (the woman) Mary. This prediction is indeed borne out as the bare noun in (25b) is judged acceptable:

(25) a. Mali gen yi ge nanhai he yi ge nühai zai jiaoshi li.
Mary with one CL boy and one CL girl at classroom inside
'Mary is in the classroom with a boy and a girl.'

b. Ta zhengzai gen **nanhai** shuohua.
3SG PROG with boy talk
'She is talking to the boy.'

Interestingly, it seems that it is even possible to have anaphoric bare nouns when the initial context is expanded incrementally to include additional participants:[15]

(26) a. Jiaoshi li zuo zhe yi ge nanhai he yi ge nühai.
Classroom inside sit PROG one CL boy and one CL girl
'A boy and a girl were sitting in the classroom.'

b. Turan yi ge xiaohai pao jin jiaoshi jiao **nanhai** gen ta chuqu.
suddenly one CL kid run in classroom ask boy with him go-out
'Suddenly, a kid ran into the classroom and asked the boy to go out with him.'

Our speculation, then, is that there is a crucial difference between the introduction of an individual in a way that implicitly shifts the initial context (as in (20c)) and the controlled introduction of an individual into a given context (as in (26)). In the first case, two distinct situations are brought into play; in the second case, a situation that satisfies uniqueness is minimally expanded to include another individual before the anaphoric noun is processed. In other words, the situation in which the last conjunct is interpreted in (26) is the same as the immediately preceding situation, namely a situation with a boy, a girl and the kid who then asks the unique boy in that same situation to go out.

We want to emphasize that we by no means think this is necessarily the last word on this issue but rather that an explanation in these terms has the right contours to capture the gradient nature of the phenomenon. As far as we can tell, it covers the cases in Jenks (2018) showing the reduced acceptability of anaphoric bare nouns and the ones we have

provided in section 2.2 showing their full acceptability.[16] *Index!*, we believe, is not the right tool for the job. Like *Maximize Presupposition* (Heim 1991), on which it is based, it leads to categorical distinctions in (un)acceptability and complementary distribution that do not characterize the anaphoric demonstratives vs. bare nouns in Mandarin.[17]

4 Conclusion

Let us consider the picture of definiteness that has emerged, expanding on the table from Jenks given earlier in table 1. There are three lexical exponents that are major players cross-linguistically: demonstratives, strong definites and weak definites, with a few languages (German, for example) reflecting this three-way distinction lexically, while many showing only a two-way distinction (English and Mandarin, for example). By introducing demonstratives into the picture we also introduce a new diagnostic environment into the calculus, the deictic use of definites, and we draw on context sentences like (15a) for determining whether a particular definite qualifies as a deictic definite:

Table 2

Definiteness marking in Mandarin, German, and English

	Mandarin	German	English
Unique definites	Bare nouns *Dem-Cl-N	Weak article *Strong article *Demonstrative	Def article *Demonstrative
Anaphoric definites	Bare nouns Dem-Cl-N	*Weak article Strong article Demonstrative	Def article Demonstrative
Contrastive deictic definites	*Bare nouns Dem-Cl-N	*Weak article Strong article Demonstrative	*Def article Demonstrative

Note that we have included the German strong article among deictic definites. This is based on Schwarz (2009: 34-37), who provides the following example:

(27) Hans ist in dem Auto [pointing at car 1] gekommen,
Hans is in the$_{strong}$ car come
nicht in dem Auto [pointing at car 2]
not in the$_{strong}$ car
'Hans came in that car, not in that car.' (Schwarz 2009:34)

Schwarz emphasizes that the strong article is not synonymous with the demonstrative because (27) is only acceptable with prosodic focus on the two definites. His discussion nevertheless underscores the point we have tried to highlight in our discussion of Mandarin, that determining whether a particular form should be aligned with a strong article definite or with a demonstrative is not straightforward.

This might also be the place to point to yet another type of distinction that needs to be included in the picture of cross-linguistic variation given in table 2. Languages can also use structural options to express definiteness. Cantonese, closely related to Mandarin, includes Classifier-Noun sequences in its inventory of definite structures. The following example, where superscripts indicate tones, shows one type of definite use:[18]

(28) 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} zing^{3}zi^{6}gaa^{1}.
Lei-sei interview PERF one CLF writer and one CLF politician
'Lei-sei interviewed a writer and a politician.'
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.' (Jenks 2018: 527)

Jenks (2018) proposes, on the basis of examples like (28), that Cantonese Classifier-Noun structures, rather than bare nouns, function canonically as anaphoric definites. Jenks' view of Cantonese bare nouns, however, is not uncontroversial (Sybesma and Sio 2008, Simpson, Soh, and Nomoto 2011, Cheng and Sybesma 2012, among others). While we

cannot go further into the facts here, the point that such sequences needs to be factored into mapping out the full cross-linguistic typology of definiteness is well-taken.[19]

Returning to table 2, we note that it raises the issue of deciding how to map a two-way lexical distinction to a three-way conceptual distinction. Highlighting the importance of this issue is part of what we hope to have achieved in our discussion of Mandarin. A second question, and one that applies to all languages, is how to capture the distribution of different types of definites within a language. If lexically distinct items were in complementary distribution, it would be clear what research strategy one should pursue. The requirement would be to define the meanings of individual items such that only one of the meanings would survive in any given context. However, the ground reality is that there are overlaps so ancilliary assumptions have to enter into anyone's account to navigate between the contexts that allow only one definite from those that allow more than one.[20]

Focusing on Mandarin, we can safely take bare NPs to function as unique definites and demonstratives to function as deictic expressions. The question that remains is the status of anaphoric definites. We have shown that both types of NPs can function in canonical anaphoric contexts. We have suggested looking at situations in terms of the individuals that define them as a way of capturing overlaps in distribution. Once we recognize overlaps, however, the question of preference enters the picture. Focusing on demonstratives vs. strong article definites, do speakers systematically accept one particular form over the other or is there individual variation on this score? Do all languages show the same preference for demonstratives vs. definites or does the

preference depend on the nature of the definite – a lexical definite as in English vs. a bare NP as in Mandarin? At this point, our knowledge of preferences related to demonstratives is not at a point where this question can be responsibly answered.

To sum up, we see Jenks' discussion of the choice between Mandarin demonstratives and bare nouns as a much needed push to expand our study of definiteness to include demonstratives and bare nouns.[21] We believe more fine-grained and sophisticated tests conducted on a statistically significant sample size are needed to empirically define the problem before a theory of competition between demonstratives and other exponents of definiteness can be clearly formulated.

References

Ahn, Dorothy. 2019. A competition-based mechanism for anaphoric expressions. Doctoral dissertation, Harvard University, Cambridge, MA.

Chao, Yuen Ren. 1968. *A grammar of spoken Chinese.* Berkeley: University of California Press.

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

Cheng, Lisa Lai-Shen, and Rint Sybesma. 1999. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30(4): 509-542.'

Cheng, Lisa Lai- Shen, and Rint Sybesma. 2012. Classifiers and DP. *Linguistic Inquiry* 43(4): 634– 650.

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

Clark, Herbert H. 1975. Bridging. In *Thinking: Readings in cognitive science*, ed. by Philip Nicholas Johnson-Laird and Peter Cathcart Wason, 411–420. Cambridge: Cambridge University Press.

Dayal, Veneeta. 2004. Number marking and (in)definiteness in kind terms. *Linguistics and Philosophy* 27(4): 393-450.

Dayal, Veneeta. 2011. Bare noun phrases. In *Semantics: An international handbook of natural language meaning*, ed. by Claudia Maeinborn, Klaus von Heusinger and Paul Portner, 2: 1088-1109. Berlin: Mouton de Gruyter.

Dayal, Veneeta. 2013. On the existential force of bare plurals across languages. In *From grammar to meaning: The spontaneous logicality of language*, ed. by Caponigro, Ivano and Carlo Cecchetto, 49-80. New York: Cambridge University Press.

Dayal, Veneeta. 2014. Bangla plural classifiers. *Language and Linguistics* 15(1): 47-87.

Dayal, Veneeta. 2016. *Questions*. New York: Oxford University Press.

Gou, Endong, Gaoqi Rao, Xiaoyue Xiao, and Jiaojiao Zang. Dashuju beijingxia BCC yuliaoku de yanzhi [The construction of the BCC Corpus in the age of Big Data]. *Yuliaoku yuyanxue* [Corpus Linguistics] (1).

Hawkins, John A. 1978. *Definiteness and indefiniteness: A study in reference and grammaticality prediction.* London: Croom Helm.

Heim, Irene. 1991. Artikel und definitheit. In *Semantics: An international handbook of contemporary research*, ed. By Arnim von Stechow and Dieter Wunderlich, 487-535. Berlin: de Gruyter.

Jenks, Peter. 2018. Articulated definiteness without articles. *Linguistic inquiry* 49:501-536.

Jiang, L. Julie. 2012. Nominal arguments and language variation. Doctoral dissertation, Harvard University, Cambridge, MA.

Jiang, L. Julie. 2020. *Nominal arguments and language variation.* New York: Oxford University Press.

Kadmon, Nirit and Fred Landman. 1993. Any, *Linguistics and Philosophy* 16: 353-422.

Kratzer, Angelika. 2007. Situations in natural language semantics, in *Stanford Encyclopedia of Philosophy Archive*, ed. by Edward N. Zalta. https://plato.stanford.edu/archives/win2017/entries/situations-semantics/.

Li, Charles N. and Sandra, A. Thompson. 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley, CA: University of California Press.

Li, Xuping. 2013. *Numeral classifiers in Chinese: The syntax-semantics interface*. Berlin: Walter de Gruyter.

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

Owusu, Augustina. 2021. Cross-categorial definiteness in Akan: A familiarity-based account. Doctoral dissertation, Rutgers University, New Brunswick, NJ.

Roberts, Craige. 2002. Demonstratives as definites, in *Information sharing: Reference and presupposition in language generation and interpretation*, ed. by Kees van Deemter and Rodger Kibble, 89-136. Stanford, CA: CSLI.

Robinson, Heather. 2005. Unexpected (in)definiteness: Plural generic expressions in Romance. Doctoral dissertation, Rutgers University, New Brunswick, NJ.

Schwarz, Florian. 2009. Two types of definites in natural language. Doctoral dissertation, University of Massachusetts, Amherst.

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

Simpson, Andrew, Hooi Ling Soh and Hiroki Nomoto. 2011. Bare classifiers and definiteness: A cross-linguistic investigation. *Studies in Language* 35(1): 168–193.

Sybesma, Rint and Joanna Ut-Seong Sio. 2008. D is for demonstrative. Investigating the position of the demonstrative in Chinese and Zhuang. *The Linguistic Review* 25: 453-478.

Trinh, Tue. 2011. Nominal reference in two classifier languages. In *Sinn und Bedeutung 15*, ed. by Ingo Reich, 629–644. Saarbrücken: Saarland University Press. http://universaar.uni-saarland.de/monographien/volltexte/2011/30/artikel/Trinh_sub15.pdf.

Wolter, Lynsey. 2006. That's that: The semantics and pragmatics of demonstrative noun phrases. Doctoral dissertation, University of California, Santa Cruz.

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

Veneeta Dayal

Department of Linguistics

Yale University

Room 208, Dow Hall, 370 Temple Street

New Haven, CT 06510

veneeta.dayal@yale.edu

Li Julie Jiang

Department of East Asian Languages and Literatures

University of Hawai‘i at Mānoa

359 Moore Hall, 1890 East-West Road

Honolulu, HI 96822

lijiang@hawaii.edu

Acknowledgments. We are extremely grateful to two anonymous reviewers for their feedback. We have also benefitted from discussions with Ming Chen, Yi-Hsun Chen, Kamil Deen, Shin Fukuda, Peter Jenks, Jess Law, Mingming Liu, Gita Martohardjono, Yimei Xiang and Beibei Xu. All remaining errors and omissions are our responsibility.

[1] The idea that there is a covert type shift in Mandarin, utilizing the same semantic operation as in the English overt definite determiner, is due to Chierchia (1998) and was adopted in Cheng and Sybesma (1999) and fully elaborated within Chierchia's system in Yang (2001). We have opted for mentioning only Yang (2001) in the text since that is the version that Jenks uses in building up his picture of Mandarin definiteness. On the interpretation of definite bare nouns as the extension of the kind, see Dayal (2011), in addition to Trinh (2011), mentioned by Jenks.

[2] Note that the condition x = *g*(1) occurs also in the scope of *iota* in (5b). This is in line with the account in Schwarz (2009). Not including this condition would have the unwelcome result that the DP would fail to refer if there were two individuals with the NP property, even though only one of them would be identical to g(1). The classic bishop or sage plant examples are relevant here.

[3] We follow the convention in Heim and Kratzer (1998, 73-76) where the material between the colon (:) and the period (.) corresponds to the presupposition and material after the period to the truth-conditional contribution: **λP**: ∃!*x*[**P**(*x*)(*s*′)] . ι*x*[**P**(*x*)(*s*′)]. Note that once the lambda expression on the left has been satisfied, the two parts are separated by the period: ∃!*x*[**student** (*x*)(*s*′)] . ι*x*[**student** (*x*)(*s*′)]. ∃!x Is to be read as *there is exactly one x such that…*

[4] The German weak and strong articles manifest themselves in the complement position of a preposition, with the former showing contraction between the preposition and the article, the latter retaining the two as independent morphemes:

(i) Hans ging zum Haus.
Hans went to- the$_{weak}$ house

(ii) Hans ging zu dem Haus.
Hans went to the$_{strong}$ house
'Hans went to the house.' Schwarz (2009: 14)

[5] An anonymous reviewer helpfully points out that our position on Mandarin definiteness resonates with ideas in Sybesma and Sio (2008). While we cannot go into the

details of that paper for reasons of space, we refer the interested reader to it, pages 466-470 in particular.

[6] See Owusu (2021) for a recent analysis of definiteness in Akan, expanding on our proposal about demonstratives and strong definites in section 3. The picture of Akan definites that emerges is much more nuanced than assumed in earlier literature.

[7] The discussion in Schwarz (2009: 52-53, 190-237, 246-253) underscores the complexity of this issue. We follow the core distinction between part-whole and producer-product types of associative anaphora that Schwarz settles on.

[8] We also checked (8a) with the bare noun *shi* in the main clause, interpreted as a definite. With this change, some speakers accepted the associated bare NP *shiren* in the second clause but some still preferred the demonstrative.

[9] Schwarz (2009: 44-49) discusses the possibility of weak article definites in anaphoric contexts and indicates how non-complementarity may arise. In discussing cases where only the weak article is possible, his example (54), Schwarz indicates how the weak article might work but does not elaborate on why the strong article is ruled out.

[10] BCC Corpus is one of the major Chinese corpora in Mainland China with diverse writing genres (newspaper, literature, *Weibo* 'microblogs', etc.), with a total number of characters around 15 billion (see Gou et al. 2016).

[11] In the case of *president*, in most contexts of use the anchor is a specific country so that the anti-uniqueness requirement rules out the demonstrative. In contexts such as a meeting of world leaders this anchoring is removed, anti-uniqueness is satisfied, and demonstratives predictably become acceptable. Similarly, if the conversation is about

several solar systems, *that/this sun* may be acceptable too. Thus there is a significant pragmatic component to the phenomenon in terms of fixing what types of contexts are evoked but the requirement itself is a presupposition.

[12] We remain neutral here on whether contextually salient entities (like the sun) are represented in the domain of discourse, since it does not affect the interpretation of demonstratives that we are concerned with.

[13] We do not identify the definite meaning of the Mandarin bare noun with that of the English definite, since there are potential technical differences due to the fact that the Mandarin bare nouns are also kind terms.

[14] See the discussion of bridging in Schwarz (2009) in part-whole cases for relevant discussion.

[15] One of our consultants finds (25)-(26) to be better than (2b, c), but not fully acceptable; another finds (2c) to be acceptable but finds the contrast between (2b) and (25)-(26) as reported here. However, it is worth noting that even the first consultant considers (11b, c) fully grammatical.

[16] The only potential outlier is (14b) but the excerpt is part of a larger text which includes Old Jolyon.

[17] The English definite and indefinite article, for example, are typically in complementary distribution once the epistemic knowledge of the participants is factored in, as predicted by *Maximize Presupposition*. The Mandarin bare noun and the demonstrative are not in complementary distribution, just as the English definite and demonstrative are not.

[18] Another language worth mentioning here is Bangla, where definiteness is marked by fronting of N to a position before the classifier (see Dayal 2014 and references there). See Jiang (2012, 2020) for more on possible variations among classifier languages.

[19] An anonymous reviewer asks how our proposed line of inquiry would apply to Cantonese. As our brief comment indicates, the empirical generalizations need more careful evaluation. We emphasize, however, that any cross-linguistic extension designed on the basis of Mandarin, ours as well as Jenks' *Index!*, would have to be parameterized if individual languages differ in their preference for demonstratives over bare NPs, as has been claimed by Ahn (2019). Our own position, as we emphasize in this conclusion, is that further theorizing has to wait till we have more data to formulate sound empirical generalizations on.

[20] See, for example, Schwarz (2009:281-286) on the overlaps in distribution of the two articles and Schwarz (2009:290-292) on the distribution of demonstratives, pronouns and definites.

[21] Ahn (2019) also looks at the distribution of bare nouns vs. demonstratives in a number of languages but her account is based on a small number of speakers. And for at least one language she notes variation among the few speakers she consulted.